

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 03:24:37 ; Search time 216 Seconds  
(without alignments)  
6219.278 Million cell updates/sec

Title: US-09-787-844-1

Perfect score: 1081

Sequence: 1 aggagcagagggggcgctca.....acattccagttgagcctgc 1081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Published\_Applications\_NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1055.8	97.7	1100	9	US-09-905-291A-256
2	1055.8	97.7	1100	9	Sequence 256, App
3	1055.8	97.7	1100	9	Sequence 256, App
4	1055.8	97.7	1100	9	Sequence 256, App
5	1055.8	97.7	1100	9	Sequence 256, App
6	1055.8	97.7	1100	9	Sequence 256, App
7	1055.8	97.7	1100	9	Sequence 256, App
8	1055.8	97.7	1100	9	Sequence 256, App
9	1055.8	97.7	1100	9	Sequence 256, App
10	1055.8	97.7	1100	9	Sequence 256, App
11	1055.8	97.7	1100	9	Sequence 256, App
12	1055.8	97.7	1100	9	Sequence 256, App
13	1055.8	97.7	1100	9	Sequence 256, App
14	1055.8	97.7	1100	9	Sequence 256, App
15	1055.8	97.7	1100	9	Sequence 256, App
16	1055.8	97.7	1100	9	Sequence 256, App
17	1055.8	97.7	1100	9	Sequence 256, App
18	1055.8	97.7	1100	9	Sequence 256, App
19	1055.8	97.7	1100	9	Sequence 256, App

20	1055.8	97.7	1100	9	US-09-904-119-256
21	1055.8	97.7	1100	9	US-09-904-956-256
22	1055.8	97.7	1100	9	US-09-907-794-256
23	1055.8	97.7	1100	9	US-09-902-692-256
24	1055.8	97.7	1100	9	US-09-903-520-256
25	1055.8	97.7	1100	9	US-09-903-943-256
26	1055.8	97.7	1100	9	US-09-904-462-256
27	1055.8	97.7	1100	9	US-09-905-056-256
28	1055.8	97.7	1100	9	US-09-907-925-256
29	1055.8	97.7	1100	9	US-09-904-553-256
30	1055.8	97.7	1100	9	US-09-903-381-256
31	1055.8	97.7	1100	9	US-09-905-064-256
32	1055.8	97.7	1100	9	US-09-905-088-256
33	1055.8	97.7	1100	9	US-09-907-575-256
34	1055.8	97.7	1100	9	US-09-902-759-256
35	1055.8	97.7	1100	9	US-09-905-075-256
36	1055.8	97.7	1100	9	US-09-902-634-256
37	1055.8	97.7	1100	9	US-09-902-713-256
38	1055.8	97.7	1100	9	US-09-907-979-256
39	1055.8	97.7	1100	10	US-09-909-320-256
40	1055.8	97.7	1100	10	US-09-909-088-256
41	1054.8	97.6	1077	10	US-09-885-441-20
42	1045.8	96.7	1082	10	US-09-885-441-19
43	493.4	45.6	529	10	US-09-833-790-374
C 44	465.4	43.1	575	9	US-10-015-219-687
C 45	465.4	43.1	575	10	US-09-777-564-687

ALIGNMENTS

RESULT 1  
US-09-905-291A-256  
; Sequence 256, Application US/09905291A  
; Patent No. US20020160374A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/905,291A  
; CURRENT FILING DATE: 2001-07-12  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222

;; PRIOR FILING DATE: 1999-07-28  
;; PRIOR APPLICATION NUMBER: PCT/US99/20594  
;; PRIOR FILING DATE: 1999-09-08  
;; PRIOR APPLICATION NUMBER: PCT/US99/20944  
;; PRIOR FILING DATE: 1999-09-13  
;; PRIOR APPLICATION NUMBER: PCT/US99/21090  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/21547  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/23089  
;; PRIOR FILING DATE: 1999-10-05  
;; PRIOR APPLICATION NUMBER: PCT/US99/28214  
;; PRIOR FILING DATE: 1999-11-29  
;; PRIOR APPLICATION NUMBER: PCT/US99/28313  
;; PRIOR FILING DATE: 1999-11-30  
;; PRIOR APPLICATION NUMBER: PCT/US99/28564  
;; PRIOR FILING DATE: 1999-12-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/28565  
;; PRIOR FILING DATE: 1999-12-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/30095  
;; PRIOR FILING DATE: 1999-12-16  
;; PRIOR APPLICATION NUMBER: PCT/US99/30911  
;; PRIOR FILING DATE: 1999-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US99/30999  
;; PRIOR FILING DATE: 1999-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US00/00219  
;; PRIOR FILING DATE: 2000-01-05  
;; NUMBER OF SEQ ID NOS: 423  
;; SEQ ID NO 256  
;; LENGTH: 1100  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-09-905-291A-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	21	GGCGGGGAGAGAGGCCATGGCGCGGGGGGCTGCTGCTGGCGGTGCTGCTGGC	80
Db	1	GGCGGGGAGAGAGGCCATGGCGCGGGGGGCTGCTGCTGGCGGTGCTGCTGGC	60
QY	81	TCGGGCTGGACTCAGAACCGGAGTCGAGAGGGGGCCCTTATCAGGACCATGCGG	140
Db	61	TCGGGCTGGACTCAGAACCGGAGTCGAGAGGGGGCCCTTATCAGGACCATGCGG	120
QY	141	CCGACGGGTATCAGCTCGCGATCGTGGGATTCGCCACGTATCGGAGTGAGCTGCTCAGCCA	260
Db	121	CCGACGGGTATCAGCTCGCGATCGTGGGATTCGCCACGTATCGGAGTGAGCTGCTCAGCCA	240
QY	201	GTGGCAGGGAGGCTCGCGCTGTGGGATTCGCCACGTATCGGAGTGAGCTGCTCAGCCA	320
Db	181	GTGGCAGGGAGGCTCGCGCTGTGGGATTCGCCACGTATCGGAGTGAGCTGCTCAGCCA	300
QY	261	CGGCTGGGACATCAGCGGGGCGACTGCTTTGAAACCTATAGTGACCTTAGTCCCTC	380
Db	241	CGGCTGGGACATCAGCGGGGCGACTGCTTTGAAACCTATAGTGACCTTAGTCCCTC	360
QY	321	CGGGTGGATGTCAGTGGGCGGAGTGCATGCTCCATGCCATCCTTCTGAGGCGCTGAGGC	440
Db	301	CGGGTGGATGTCAGTGGGCGGAGTGCATGCTCCATGCCATCCTTCTGAGGCGCTGAGGC	420
QY	381	CTACTACACCCGTTACTTGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC	500
Db	361	CTACTACACCCGTTACTTGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC	480
QY	441	ACCCTATGACATGCTTGGTGAAGCTGTGTGCACCTGTACCTACACTAACACATCCA	560
Db	421	ACCCTATGACATGCTTGGTGAAGCTGTGTGCACCTGTACCTACACTAACACATCCA	540
QY	501	GCCCATCTGTCTCCAGGCGCTCCACATTTGAGTTTGAACCGGACAGACTGCTGGGTGAC	
Db	481	GCCCATCTGTCTCCAGGCGCTCCACATTTGAGTTTGAACCGGACAGACTGCTGGGTGAC	

QY	561	TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCCCTCCAGGAAGT	620
Db	541	TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCCCTCCAGGAAGT	600
QY	621	TCAGTGCCTCATATAAACAACCTCTATGTGCAACACCCCTCTCTCTCAAGTACAGTTTCGG	680
Db	601	TCAGTGCCTCATATAAACAACCTCTATGTGCAACACCCCTCTCTCTCAAGTACAGTTTCGG	660
QY	681	CAAGGACATCTTTGGAGACATGGTTTGTCTGGCAATGCCAAGCGGGAAGGATGCGCTG	740
Db	661	CAAGGACATCTTTGGAGACATGGTTTGTCTGGCAATGCCAAGCGGGAAGGATGCGCTG	720
QY	741	CTTCGGTACTCAGGTGGACCCCTTGGGCTGTAAACAAGATGGACTGTGTATCAGATTGG	800
Db	721	CTTCGGTACTCAGGTGGACCCCTTGGGCTGTAAACAAGATGGACTGTGTATCAGATTGG	780
QY	801	AGTCGTGAGCTGGGGAGTGGGCTGTGTCGGCCCAATCGGCCCGGCTGTCTACCAATAT	860
Db	781	AGTCGTGAGCTGGGGAGTGGGCTGTGTCGGCCCAATCGGCCCGGCTGTCTACCAATAT	840
QY	861	CAGCCACCATCTTTGAGTGGATCCAGAAAGCTGATGGCCAGAGTGGCATGTCCAGCCAGA	920
Db	841	CAGCCACCATCTTTGAGTGGATCCAGAAAGCTGATGGCCAGAGTGGCATGTCCAGCCAGA	900
QY	921	CCCTCTCTGGCCACTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	980
Db	901	CCCTCTCTGGCCACTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	960
QY	981	CTGAGCCTACTCTGAGCCCATGAGCCTGGGCGGCACTGCCAAGTCCAGTCTCTCTCT	1040
Db	961	CTGAGCCTACTCTGAGCCCATGAGCCTGGGCGGCACTGCCAAGTCCAGTCTCTCTCT	1020
QY	1041	CTCTCTTTTGGTAAATAACACATTCAGTTGATGCT	1079
Db	1021	CTCTCTTTTGGTAAATAACACATTCAGTTGATGCT	1059

## RESULT 2

US-09-902-853-256  
; Sequence 256, Application US/09902853  
; Publication No. US20020192659A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/902,853  
; CURRENT FILING DATE: 2001-07-10



APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/907,824  
CURRENT FILING DATE: 2001-07-17  
PRIOR APPLICATION NUMBER: 09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 256  
LENGTH: 1100  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-09-907-824-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	21	GGCCGCGGAGAGGAGCGCATGGCGCGCGCGCGCGCTGCTGCTGGCGCTGCTGCTGGC	80
Db	1	GGCCGCGGAGAGGAGCGCATGGCGCGCGCGCGCGCTGCTGCTGGCGCTGCTGCTGGC	60
QY	81	TCGGGCTGACCTCAGGAAGCGGAGTCGCGAGGAGCGCGCGCTTATCAGGACCATGCGG	140
Db	61	TCGGGCTGACCTCAGGAAGCGGAGTCGCGAGGAGCGCGCGCTTATCAGGACCATGCGG	120
QY	141	CCGACGGGTATCAGCTCGCGCATCGTGGTGGAGAGGACGCGCACTCGGCGGTGGCC	200
Db	121	CCGACGGGTATCAGCTCGCGCATCGTGGTGGAGAGGACGCGCACTCGGCGGTGGCC	180
QY	201	GTGGCAGGAGGAGCGCTGCGGCTCGGATGCCACGTATCGGAGTGAGCGCTCTCAGCCA	260
Db	181	GTGGCAGGAGGAGCGCTGCGGCTCGGATGCCACGTATCGGAGTGAGCGCTCTCAGCCA	240
QY	261	CGGCTGGGACATCAGCGGGCGCACTGCTTTGAACCTATAGTAGCTAGTATCCCTC	320
Db	241	CGGCTGGGACATCAGCGGGCGCACTGCTTTGAACCTATAGTAGCTAGTATCCCTC	300

QY	321	CGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCATCTTCTTGAGCCTGCAGGC	380
Db	301	CGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCATCTTCTTGAGCCTGCAGGC	360
QY	381	CTACTACACCCGTTACTTCGTTATCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC	440
Db	361	CTACTACACCCGTTACTTCGTTATCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC	420
QY	441	ACCCCTATGACATTTGGCTTTGGTGAAGCTGCTGCACTCTGCACTACACTAAACACATCCA	500
Db	421	ACCCCTATGACATTTGGCTTTGGTGAAGCTGCTGCACTCTGCACTACACTAAACACATCCA	480
QY	501	GCCCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC	560
Db	481	GCCCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC	540
QY	561	TGGCTGGGGTACATCAAGAGGATGAGGAGTGGCCATCTCCCAACACCCCTCCAGGAAT	620
Db	541	TGGCTGGGGTACATCAAGAGGATGAGGAGTGGCCATCTCCCAACACCCCTCCAGGAAT	600
QY	621	TCAGGTGCCATCATAAACAACTCTATGTGCAACACCTCTTCTCAAGTACAGTTTCCG	680
Db	601	TCAGGTGCCATCATAAACAACTCTATGTGCAACACCTCTTCTCAAGTACAGTTTCCG	660
QY	581	CAAGGACATCTTTGGAGACATGGTTTGTGTGGCAATGCCCAAGCGGGAAGGATGCTG	740
Db	661	CAAGGACATCTTTGGAGACATGGTTTGTGTGGCAATGCCCAAGCGGGAAGGATGCTG	720
QY	741	CTTCGGTGACTCAGGTGGACCCCTTGGCCTGTAAACAAGATGGACTGTGATCAGATTGG	800
Db	721	CTTCGGTGACTCAGGTGGACCCCTTGGCCTGTAAACAAGATGGACTGTGATCAGATTGG	780
QY	801	AGTCGTGAGCTGGGGAGTGGGCTGTGTGCGGCCCAATCGGCCGCTGCTACACCAATAT	860
Db	781	AGTCGTGAGCTGGGGAGTGGGCTGTGTGCGGCCCAATCGGCCGCTGCTACACCAATAT	840
QY	861	CAGCCACCTTTGAGTGGATCCAGAACTGATGGCCAGAGTGGCATGTCCCAGCCAGA	920
Db	841	CAGCCACCTTTGAGTGGATCCAGAACTGATGGCCAGAGTGGCATGTCCCAGCCAGA	900
QY	921	CCCTCTGCGCCACTACTCTTTTCCCTCTTCTCTGCTGGGCTCTCCCACTCTGGGGCGGT	980
Db	901	CCCTCTGCGCCACTACTCTTTTCCCTCTTCTCTGCTGGGCTCTCCCACTCTGGGGCGGT	960
QY	981	CTGAGCTTACTGAGCCCATGAGCTGGGGCCACTGCGCAAGTCAGGCCCTGGTCTCTTT	1040
Db	961	CTGAGCTTACTGAGCCCATGAGCTGGGGCCACTGCGCAAGTCAGGCCCTGGTCTCTTT	1020
QY	1041	CTGCTCTTTTGGTAAATAACACATTTCCAGTTGATGCTT	1079
Db	1021	CTGCTCTTTTGGTAAATAACACATTTCCAGTTGATGCTT	1059

## RESULT 4

US-09-907-841-256  
; Sequence 256, Application US/09907841  
; Publication No. US20020198366A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.



APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/907,841  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 256  
LENGTH: 1100  
TYPE: DNA  
ORGANISM: Homo Saplen  
US-09-907-841-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 GGCGCGGGAGAGAGAGCCATGGCGCGCGGGCGCGCTGCTGCTGGCGCTGCTGCTGCG 80  
DB 1 GGCGCGGGAGAGAGAGCCATGGCGCGCGGGCGCGCTGCTGCTGGCGCTGCTGCTGCG 60  
QY 81 TCGGGCTGGACTCAGGAAGCGGAGTCGCGAGGAGGCGGCGGCGGCGGCGGCGGCGG 140  
DB 61 TCGGGCTGGACTCAGGAAGCGGAGTCGCGAGGAGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 141 CCAGCGGCTCATCAGTCGCGCATCGGTCGGTGGAGAGAGCGCCAACTCGGGCGCTGCGC 200  
DB 121 CCAGCGGCTCATCAGTCGCGCATCGGTCGGTGGAGAGAGCGCCAACTCGGGCGCTGCGC 180  
QY 201 GTGCGAGGGAGCGCTGCGCGCTGTGGGATCCACGATGCGGAGTGGAGCTGCTGCGAGCCA 260  
DB 181 GTGCGAGGGAGCGCTGCGCGCTGTGGGATCCACGATGCGGAGTGGAGCTGCTGCGAGCCA 240  
QY 261 CCGTGGGCACTACGCGGGGCGACTGCTTTGAAACCTATAGTACCTTAGTATGATCCCTC 320  
DB 241 CCGTGGGCACTACGCGGGGCGACTGCTTTGAAACCTATAGTACCTTAGTATGATCCCTC 300  
QY 321 CCGGTGATGGTCAGTTGGCCAGCTGACTTCCATGCCATCCTCTTGGAGCCTGCGAGGC 380  
DB 301 CCGGTGATGGTCAGTTGGCCAGCTGACTTCCATGCCATCCTCTTGGAGCCTGCGAGGC 360  
QY 381 CTAACACCGGTTACTTCGTATCGAATATCTATCTAGCGCCCTCGCTACCTGGGGAATTC 440

DB 361 CTACTACACCGTTACTTCGTATCGAATATCTATCTAGCGCCCTCGCTACCTGGGAATTC 420  
QY 441 ACCCTATGACATTCGCTTGGTGAAGCTGCTGACCTGTGACCTACACTAACACATCCA 500  
DB 421 ACCCTATGACATTCGCTTGGTGAAGCTGCTGACCTGTGACCTACACTAACACATCCA 480  
QY 501 GCCCATCTGCTCCAGGCGCTCCACATTTGAGTTGAGAACCGGAGAGTCTGCTGGGTGAC 560  
DB 481 GCCCATCTGCTCCAGGCGCTCCACATTTGAGTTGAGAACCGGAGAGTCTGCTGGGTGAC 540  
QY 561 TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCGCATCTCCACACACCTCCAGGAAT 620  
DB 541 TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCGCATCTCCACACACCTCCAGGAAT 600  
QY 621 TCAGGTGCGCATATAAACAACACTCTATGTGCAACACCTCTTCTCTCAAGTACAGTTCCG 680  
DB 601 TCAGGTGCGCATATAAACAACACTCTATGTGCAACACCTCTTCTCTCAAGTACAGTTCCG 660  
QY 681 CAAGSACATCTTGGAGACATGTTGCTGGCAATGCCAAGCGGGAAGGATGCGCTG 740  
DB 661 CAAGSACATCTTGGAGACATGTTGCTGGCAATGCCAAGCGGGAAGGATGCGCTG 720  
QY 741 CTTCCGCTGACTCAGGTGAGCCCTTGGCCTGTAAACAAGATGGACTGTGTATCAGATTGG 800  
DB 721 CTTCCGCTGACTCAGGTGAGCCCTTGGCCTGTAAACAAGATGGACTGTGTATCAGATTGG 780  
QY 801 AGTCGTGAGCTGGGAGTGGCTGTGGTGGCCCAATGCGGCCGCTGCTACACCAATAT 860  
DB 781 AGTCGTGAGCTGGGAGTGGCTGTGGTGGCCCAATGCGGCCGCTGCTACACCAATAT 840  
QY 861 CAGCACCATCTTGGAGTGGATCCAGAGCTGATGCGCCAGAGTGGCATGCTCCAGCCAGA 920  
DB 841 CAGCACCATCTTGGAGTGGATCCAGAGCTGATGCGCCAGAGTGGCATGCTCCAGCCAGA 900  
QY 921 CCCCTCTGCGCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGCGCGT 980  
DB 901 CCCCTCTGCGCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGCGCGT 960  
QY 981 CTGAGCCTACCTGAGCCCATGCGCCCTGGGCGCACTGCGCAAGTCAAGGCGCTTCTT 1040  
DB 961 CTGAGCCTACCTGAGCCCATGCGCCCTGGGCGCACTGCGCAAGTCAAGGCGCTTCTT 1020  
QY 1041 CTGCTCTGTTGGTAATAACACATTCAGTTGATGCT 1079  
DB 1021 CTGCTCTGTTGGTAATAACACATTCAGTTGATGCT 1059

RESULT 5  
US-09-904-011-256  
Sequence 256, Application US/09904011  
Publication No. US20030003530A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.





APPLICANT: Botstein, David  
APPLICANT: Denoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijawin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/906,838  
CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 256  
LENGTH: 1100  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-09-906-838-236

Matches 1057; Conservative 0; Mismatches 2; Indels 0; Caps 0;  
QY 21 GCGCGGGAGAGAGAGCCATGGCGCGCGCGGGGCGCTGCTGCTGCGCTGCTGCTGCG 80  
DB 1 GCGCGGGAGAGAGAGCCATGGCGCGCGCGGGGCGCTGCTGCTGCGCTGCTGCTGCG 60  
QY 81 TCGGGCTGGACTCAGGAAGCCGGAGTCGCGAGGCGCGCGCCCTTATCAGGACCATCGG 140  
DB 61 TCGGGCTGGACTCAGGAAGCCGGAGTCGCGAGGCGCGCGCCCTTATCAGGACCATCGG 120  
QY 141 CCGACGGGTATCATCGCTCGCGCATCGTGGGTGGAGAGAGCGCGAACTCGGGGCTGGCC 200  
DB 121 CCGACGGGTATCATCGCTCGCGCATCGTGGGTGGAGAGAGCGCGAACTCGGGGCTGGCC 180  
QY 201 GTGGAGGGGAGCTGCGCGCTGCGGATTCACAGTATCCACGATGCGGAGTGGAGCTGCTCAGCCA 260  
DB 181 GTGGAGGGGAGCTGCGCGCTGCGGATTCACAGTATCCACGATGCGGAGTGGAGCTGCTCAGCCA 240  
QY 261 CCGCTGGGCACTCACGGCGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTC 320  
DB 241 CCGCTGGGCACTCACGGCGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTC 300  
QY 321 CCGGTGGATGGTCCAGTTGGCGAGCTGACTTCCATGCCATCTCTGGAGCCTGCGAGC 380  
DB 301 CCGGTGGATGGTCCAGTTGGCGAGCTGACTTCCATGCCATCTCTGGAGCCTGCGAGC 360  
QY 381 CTACTACACCGCTTACTTCTGATATCTATCTGAGCCCTGCTGCTGCTGCTGCTGCTGCT 440  
DB 361 CTACTACACCGCTTACTTCTGATATCTATCTGAGCCCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 441 ACCCTATGACATTCGCTTGGTGAAGCTGTGTCACCTGTGTCACCTGTGTCACCTGTGTCAC 500  
DB 421 ACCCTATGACATTCGCTTGGTGAAGCTGTGTCACCTGTGTCACCTGTGTCACCTGTGTCAC 480  
QY 501 GCGCATCTCTCCAGCGCTCCACATTTGAGTTTGAAGCCGAGACAGCTGCTGGGTGAC 560  
DB 481 GCGCATCTCTCCAGCGCTCCACATTTGAGTTTGAAGCCGAGACAGCTGCTGGGTGAC 540  
QY 561 TGGCTGGGGGTACATCAAGAGAGATGAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620  
DB 541 TGGCTGGGGGTACATCAAGAGAGATGAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 621 TCAGTTCGCCATCAATAACAATCTATGTGCAACCACTCTTCTCTCAAGTACAGTTCCG 680  
DB 601 TCAGTTCGCCATCAATAACAATCTATGTGCAACCACTCTTCTCTCAAGTACAGTTCCG 660  
QY 681 CAAGGACATCTTTGGAGACATGTTTGTCTGCTGCAATGCGGAGCGGAGGATGCTGCTG 740  
DB 661 CAAGGACATCTTTGGAGACATGTTTGTCTGCTGCAATGCGGAGCGGAGGATGCTGCTG 720  
QY 741 CTTGGTGAATCAGGTGGACCTTTGGCTGTAAACAGAAATGGACTGTGATCAGATTGG 800  
DB 721 CTTGGTGAATCAGGTGGACCTTTGGCTGTAAACAGAAATGGACTGTGATCAGATTGG 780  
QY 801 ACTGTGAGCTGGGAGTGGGTGCTGGTGGCGCCCAATCGGCGGCTGCTACACCAATAT 860  
DB 781 ACTGTGAGCTGGGAGTGGGTGCTGGTGGCGCCCAATCGGCGGCTGCTACACCAATAT 840  
QY 861 CAGCCACCACTTTGAGTGGATCCAGAGCTGATGCGCCAGAGTGGCATGCTCCAGCCAGA 920  
DB 841 CAGCCACCACTTTGAGTGGATCCAGAGCTGATGCGCCAGAGTGGCATGCTCCAGCCAGA 900  
QY 921 CCCCTCCTGGCCACTACTCTTTTCCCTTCTCTGCGGCTCTCCCACTCTCTGGGCGGCT 980  
DB 901 CCCCTCCTGGCCACTACTCTTTTCCCTTCTCTGCGGCTCTCTGCGGCTCTCTGGGCGGCT 960  
QY 981 CTGAGCCTACCTGAGCCCATGAGCCTTGGGCGCACTGCGCAAGTCAAGGCGCTGCTTCTT 1040  
DB 961 CTGAGCCTACCTGAGCCCATGAGCCTTGGGCGCACTGCGCAAGTCAAGGCGCTGCTTCTT 1020  
QY 1041 CTGCTCTCTTGGTAAACACATTCAGTTGATGCCT 1079  
DB 1021 CTGCTCTTGGTAAACACATTCAGTTGATGCCT 1059





Db 841 CAGCCACCACTTGTAGTGGATCCAGAGCTGATGGCCAGAGTGCATGTCTCCAGCCAGA 900  
QY 921 CCCCTCCGCGCACTACTCTTTCCCTCTCTCTGGGCTCTCCACACTCCCTGGGCGCGT 980  
Db 901 CCCCTCCGCGCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACACTCCCTGGGCGCGT 960  
QY 981 CTGAGCTTACCTGAGCCCATGAGCTGGGCGCACTGCAAGTCAAGGCTCTGGTCTCTCT 1040  
Db 961 CTGAGCTTACCTGAGCCCATGAGCTGGGCGCACTGCAAGTCAAGGCTCTGGTCTCTCT 1020  
QY 1041 CTGCTCTTGGTGAATTAACACATTCAGTTGATGCCT 1079  
Db 1021 CTGCTCTTGGTGAATTAACACATTCAGTTGATGCCT 1059

RESULT 10

US-09-904-820-256

; Sequence 256, Application US/09904820  
; Publication No. US20030036094A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Flivaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/904,820

; CURRENT FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: 09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 256  
; LENGTH: 1100  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-904-820-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 GGCCCGGGAGAGAGGCGCATGGCGCGCGGGCGGCTGCTGCTGGCGCTGCTGCTGGC 80  
Db 1 GGCCCGGGAGAGAGGCGCATGGCGCGCGGGCGGCTGCTGCTGGCGCTGCTGCTGGC 60  
QY 81 TGGGCTGGACTCAGGAAGCCGAGTCCAGGAGCGGCGGCGCTTATCAGGACCATGGG 140  
Db 61 TCGGCTGGACTCAGGAAGCCGAGTCCAGGAGCGGCGGCGCTTATCAGGACCATGGG 120  
QY 141 CCGACGGGTATCAGCTGCGCGCATCGTGGTGGAGAGGACGCCGACCTGGGCGTGGCC 200  
Db 121 CCGACGGGTATCAGCTGCGCGCATCGTGGTGGAGAGGACGCCGACCTGGGCGTGGCC 180  
QY 201 GTGGCAGGGGAGCGCTGCGCGCTGTGGGATTCCACACCTATGCGGAGTGAGCTCTCAGCCA 260  
Db 181 GTGGCAGGGGAGCGCTGCGCGCTGTGGGATTCCACACCTATGCGGAGTGAGCTCTCAGCCA 240  
QY 261 CCGTGGGCACTCAGCGGGGCGCACTGCTTTGAAACCTATATAGTACCTTAGTGATCCCTC 320  
Db 241 CCGTGGGCACTCAGCGGGGCGCACTGCTTTGAAACCTATATAGTACCTTAGTGATCCCTC 300  
QY 321 CCGGTGGATGTCACAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTTGAGAGCTCAGCG 380  
Db 301 CCGGTGGATGTCACAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTTGAGAGCTCAGCG 360  
QY 381 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 440  
Db 361 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 420  
QY 441 ACCCTATGACATTCGCTTGGTGAAGCTGTGTCACCTGTGTCACCTACACTAAACATCA 500  
Db 421 ACCCTATGACATTCGCTTGGTGAAGCTGTGTCACCTGTGTCACCTACACTAAACATCA 480  
QY 501 GCCCATCTCTCCAGGCGCTCCACATTTGAGTTTGGAGACCGAGAGTCTGGGTGAC 560  
Db 481 GCCCATCTCTCCAGGCGCTCCACATTTGAGTTTGGAGACCGAGAGTCTGGGTGAC 540  
QY 561 TGGCTGGGGGTACATCAAGAGAGTGAAGCACTGCGCATCTCCACACACCTCCAGGAAGT 620  
Db 541 TGGCTGGGGGTACATCAAGAGAGTGAAGCACTGCGCATCTCCACACACCTCCAGGAAGT 600  
QY 621 TCAGTGCCTCATATAAACAACCTATATGCAACCACTCTTCTTCAAGTACAGTTTCG 680  
Db 601 TCAGTGCCTCATATAAACAACCTATATGCAACCACTCTTCTTCAAGTACAGTTTCG 660  
QY 681 CAAGGACATCTTTGGAGACATGGTTTGTCTGGCAATCCCCAAGGCGGGAAGATCCCG 740  
Db 661 CAAGGACATCTTTGGAGACATGGTTTGTCTGGCAATCCCCAAGGCGGGAAGATCCCG 720









Db 841 CAGCCACCATTGAGTGGATCCAGAGCTGATGCGCCAGAGTGCATGTCCAGCCAGA 900  
QY 921 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGCCGGT 980  
Db 901 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGCCGGT 960  
QY 981 CTGAGCTTACCTGAGCCCATGAGCCTGGGGCCACTGCAAGTCAAGGCCCTGGTTCTCTT 1040  
Db 961 CTGAGCTTACCTGAGCCCATGAGCCTGGGGCCACTGCAAGTCAAGGCCCTGGTTCTCTT 1020  
QY 1041 CTGCTCTTCTGGTAATAAACACATCCAGTTGATGCT 1079  
Db 1021 CTGCTCTTCTGGTAATAAACACATCCAGTTGATGCT 1059

RESULT 14

US-09-906-646-256  
; Sequence 256, Application US/09906646  
; Publication No. US20030039971A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/906.646  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 256  
; LENGTH: 1100  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; US-09-906-646-256  
  
Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 21 GGGCCGGGAGAGAGGCCATGGCGCGCGGGCGCTGTCTGTGGCGCTCTCTGTGCG 80  
Db 1 GGGCCGGGAGAGAGGCCATGGCGCGCGGGCGCTGTCTGTGGCGCTCTCTGTGCG 60  
QY 81 TCGGCTGGACTCAGGAAGCCGGAGTCCAGGAGGCGGCGCCCTTATCAGGACCATGGG 140  
Db 61 TCGGCTGGACTCAGGAAGCCGGAGTCCAGGAGGCGGCGCCCTTATCAGGACCATGGG 120  
QY 141 CCGACGGGTTCATCAGCTCGCGCATCTGGGTGGAGAGACGCCGCACTCGGGCGCTTGGCC 200  
Db 121 CCGACGGGTTCATCAGCTCGCGCATCTGGGTGGAGAGACGCCGCACTCGGGCGCTTGGCC 180  
QY 201 GTGGCAGGGAGCGCTGTGGGATTTCCACATATGTCGGGAGTGAGCGCTCTCAGCCA 260  
Db 181 GTGGCAGGGAGCGCTGTGGGATTTCCACATATGTCGGGAGTGAGCGCTCTCAGCCA 240  
QY 261 CCGCTGGGCACTCAGCGGCGGCGCTGCTTGAACCTATAGTGACCTTAGTGATCCCTC 320  
Db 241 CCGCTGGGCACTCAGCGGCGGCGCTGCTTGAACCTATAGTGACCTTAGTGATCCCTC 300  
QY 321 CCGGTGGATGTGTCAGTTTGGCCAGCTGACCTTCCATGCCATCCTTCTTGAGCGCTCAGCG 380  
Db 301 CCGGTGGATGTGTCAGTTTGGCCAGCTGACCTTCCATGCCATCCTTCTTGAGCGCTCAGCG 360  
QY 381 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 440  
Db 361 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 420  
QY 441 ACCCTATGACATTGCTTGGTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCA 500  
Db 421 ACCCTATGACATTGCTTGGTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCA 480  
QY 501 GCCCATCTGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC 560  
Db 481 GCCCATCTGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC 540  
QY 561 TGGCTGGGGGTACATCAAGAGGATGAGGCATGTCATCTCCCCACACCTCCAGGAAGT 620  
Db 541 TGGCTGGGGGTACATCAAGAGGATGAGGCATGTCATCTCCCCACACCTCCAGGAAGT 600  
QY 621 TCAGTCCGCCATCAAAACAACTCTATGTGAACCACTCTTCTTCAAGTACAGTTTCG 680  
Db 601 TCAGTCCGCCATCAAAACAACTCTATGTGAACCACTCTTCTTCAAGTACAGTTTCG 660  
QY 681 CAAGGACATCTTTGGAGACATGTTTGTGGCAATGCCCAAGCGGAGGATGCGCTG 740  
Db 661 CAAGGACATCTTTGGAGACATGTTTGTGGCAATGCCCAAGCGGAGGATGCGCTG 720  
QY 741 CTTCGCTGACTCAGGTGGACCTTGGCCCTGTAAACAAGATGACCTGTGTATCAGATTGG 800

Db	721	CTTCGGTGACTCAGGTGGACCCCTTGCCCTGTAAACAAGAAATGCACCTGTGGTATCATGATGG	780
QY	801	AGTCGTGAGCTGGGAGTGGGTGTGGTCGGCCCAATCGCCCGGTGCTACACCAATAT	860
Db	781	AGTCGTGAGCTGGGAGTGGGTGTGGTCGGCCCAATCGCCCGGTGCTACACCAATAT	840
QY	861	CAGCCACCACCTTGAGTGGATCCAGAGCTGATGCCCAGAGTGGCATGTCCCAGCCAGA	920
Db	841	CAGCCACCACCTTGAGTGGATCCAGAGCTGATGCCCAGAGTGGCATGTCCCAGCCAGA	900
QY	921	CCCTCCTGGCCACTACTCTTTTCCCCTTCTCTGCGCTCTCCCACCTGGGGCCGGT	980
Db	901	CCCTCCTGGCCACTACTCTTTTCCCCTTCTCTGCGCTCTCCCACCTGGGGCCGGT	960
QY	981	CTGAGCCTACCTTGAGCCCATGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCCTTT	1040
Db	961	CTGAGCCTACCTTGAGCCCATGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCCTTT	1020
QY	1041	CTGTCTGTGGTGAATAACACATATCCAGTTGATGCT	1079
Db	1021	CTGTCTGTGGTGAATAACACATATCCAGTTGATGCT	1059

## RESULT 15

US-09-906-700-256

Sequence 256, Application US/09906700

Publication No. US20030039972A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, A.

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/906,700

CURRENT FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,222

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/20594

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/21090

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547

```

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/230899
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 256
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: Homo Sapien
US-99-906-700-256

```

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%;  
Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0;

QY	21	GGCGCGGGAGAGAGGCCATGGCGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGCG	80
Db	1	GGCGCGGGAGAGAGGCCATGGCGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGCG	60
QY	81	TCGGCTGGACTCAGAAAGCCGAGTGCAGAGGGCGGCCCTTATCAGGACCATGCGG	140
Db	61	TCGGCTGGACTCAGAAAGCCGAGTGCAGAGGGCGGCCCTTATCAGGACCATGCGG	120
QY	141	CCGACGGGTATCACGTCGCGCATCTGGTGGAGAGGACGCCAACTCGGGCGTTGGCC	200
Db	121	CCGACGGGTATCACGTCGCGCATCTGGTGGAGAGGACGCCAACTCGGGCGTTGGCC	180
QY	201	GTGGCAGGAGGCGCTGCGCCTGTGGATCCACATATGCGGAGTGAACCTGCTCAGCCA	260
Db	181	GTGGCAGGAGGCGCTGCGCCTGTGGATCCACATATGCGGAGTGAACCTGCTCAGCCA	240
QY	261	CGCTGGGCACACGCGGGCGCACTGCTTTGAAACCTATAGTGACCTTACTGATCCCTC	320
Db	241	CGCTGGGCACACGCGGGCGCACTGCTTTGAAACCTATAGTGACCTTACTGATCCCTC	300
QY	321	CGGCTGGATGTCAGTTTGGCCAGCTGACATTCATGCCATCTCTGGAGCCTGCAGGC	380
Db	301	CGGCTGGATGTCAGTTTGGCCAGCTGACATTCATGCCATCTCTGGAGCCTGCAGGC	360
QY	381	CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATC	440
Db	361	CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATC	420
QY	441	ACCTTACACATTCGCTTGGTGAAGCTGTGACACCTGTCACTTACACTAAACACATCCA	500
Db	421	ACCTTACACATTCGCTTGGTGAAGCTGTGACACCTGTCACTTACACTAAACACATCCA	480
QY	501	GCCATCTGCTCTCAGGCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC	560
Db	481	GCCATCTGCTCTCAGGCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC	540
QY	561	TGGCTGGGGTACATCAAGAGGATGAGGCACCTGCCATCTCCCCACACCCCTCCAGGAAGT	620
Db	541	TGGCTGGGGTACATCAAGAGGATGAGGCACCTGCCATCTCCCCACACCCCTCCAGGAAGT	600
QY	621	TCAGGTCCGCATCATAAACAACTCTATGTGCACACACCTTTCCTCAAGTACAGTTTCGG	680
Db	601	TCAGGTCCGCATCATAAACAACTCTATGTGCACACACCTTTCCTCAAGTACAGTTTCGG	660

Qy	681	CAAGGACATCTTTGGAGACATGGTTTGCTGGCAATGCCAAGGCGGGAAGGATGCCTG	740
Db	661	CAAGGACATCTTTGGAGACATGGTTTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTG	720
Qy	741	CTTCGGTGACTCAGGTGACCCCTTGGCCCTGTAAACAAGAAATGGACTGTGGTATCAGATTGG	800
Db	721	CTTCGGTGACTCAGGTGACCCCTTGGCCCTGTAAACAAGAAATGGACTGTGGTATCAGATTGG	780
Qy	801	AGTCGTGAGCTGGGGAGTGGGCTGTGTCGGCCCAATCGGCCCGGTGTCTACACCAATAT	860
Db	781	AGTCGTGAGCTGGGGAGTGGGCTGTGTCGGCCCAATCGGCCCGGTGTCTACACCAATAT	840
Qy	861	CAGCCACCACCTTTGAGTGGATCCAGAAGCTGATGGCCCGAGAGTGGCATGTCCAGCCAGA	920
Db	841	CAGCCACCACCTTTGAGTGGATCCAGAAGCTGATGGCCCGAGAGTGGCATGTCCAGCCAGA	900
Qy	921	CCCTTCCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCCTGGGGCCGGT	980
Db	901	CCCTTCCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCCTGGGGCCGGT	960
Qy	981	CTGAGCCTACCTGAGCCCATGAGCCTGGGGCCACTGCCAAGTCAAGGCCCTGGTTCTCTT	1040
Db	961	CTGAGCCTACCTGAGCCCATGAGCCTGGGGCCACTGCCAAGTCAAGGCCCTGGTTCTCTT	1020
Qy	1041	CTGCTCTTCTTTGGTAATAAACACATTCAGTTGATGCCT	1079
Db	1021	CTGCTCTTCTTTGGTAATAAACACATTCAGTTGATGCCT	1059

Search completed: May 10, 2003, 05:00:59  
Job time : 219 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 03:18:27 ; Search time 2077 seconds

(without alignments)

8429.139 Million cell updates/sec

Title: 05-09-787-844-1

Perfect score: 1081

Sequence: 1 aggagcagggggcgctca.....acattccagttgacgtcgc 1081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	906.8	83.9	936	9 AL578261	AL578261 AL578261
2	862.4	79.8	876	9 AL555870	AL555870 AL555870
3	777.2	71.9	871	13 BI334059	BI334059 602997320
4	760.4	70.3	777	13 BI259237	BI259237 602970095
5	733.2	67.8	878	13 BI827608	BI827608 603073575
c 6	624.8	57.8	633	9 AI126185	AI126185 qd81h05.x

7	567.6	52.5	705	13	BM019026
8	551.4	51.0	682	13	BM017444
9	549.8	50.9	1219	13	BM553333
10	523.4	48.4	824	13	BM018961
c 11	523	48.4	591	9	AA161187
12	517.4	47.9	1286	14	BM911057
13	501.6	46.0	825	13	BI257649
14	486.2	45.4	1050	11	AK006271
15	483.6	44.7	624	13	BM018436
16	478	44.2	625	12	EG717268
17	474	43.8	628	12	EG717347
18	468	43.3	479	9	AA448095
19	452	41.8	484	14	BQ252128
c 20	451.4	41.4	514	9	AA620757
21	447.4	41.4	510	14	BM759973
c 22	441.6	40.9	462	9	AI016399
c 23	412.2	38.1	487	9	AA525010
c 24	383	35.4	399	9	AI278269
c 25	383	35.4	412	9	AA448166
c 26	381.4	35.3	402	9	AA729680
c 27	379.8	35.1	406	9	AI150137
c 28	361.4	33.4	383	9	AA447596
c 29	351.2	32.5	454	13	BI257498
30	337	31.2	1089	11	AK016509
c 31	326.4	30.2	411	9	AA149942
c 32	318.2	29.4	373	9	AA401115
c 33	311.2	28.8	331	9	AI217989
c 34	306.4	28.3	334	9	AA552523
c 35	306.2	28.3	475	10	AW971363
c 36	295.2	27.3	542	9	AI326140
c 37	288.4	26.7	351	9	AA860356
c 38	269	24.9	950	10	BE546936
c 39	236.2	21.9	359	9	AI926900
c 40	223.6	20.7	582	13	BG987202
c 41	222	20.5	482	12	BF405556
c 42	222	20.5	624	12	BF404444
c 43	222	20.5	625	12	BF403421
c 44	222	20.5	690	12	BF405109
c 45	212	19.6	431	12	BF405483

#### ALIGNMENTS

RESULT 1  
LOCUS AL578261/c  
DEFINITION AL578261 LTI\_NFL006\_PL2 Homo sapiens cdna clone CS0DK002YM07 3 prime, mRNA sequence.  
ACCESSION AL578261  
VERSION AL578261.1 GI:12942167  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
AUTHORS 1 (bases 1 to 936)  
TITLE Full-length cdna libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. .936  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DK002YM07"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cdna was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com

BASE COUNT 206 a 256 c 288 g 180 t 6 others  
ORIGIN

Query Match 83.9%; Score 906.8; DB 9; Length 936;  
Best Local Similarity 98.5%; Pred. No. 1.3e-223;  
Matches 909; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

```
QY 107 CGCAGAGCGCGCCCTTATCAGGACCATGCGCGCGAGCGGTATCATCGTCCGCGCATCG 166
|||||
DB 936 CGCAGAGCGCGCGCTTATCAGGACCATGCGCGCGAGCGGTATCATCGTCCGCGCATCG 877

QY 167 TGGGTGGAGAGCAGCGCAACTCGGGCGTGGCGGTGGCGAGGGAGCGCTGCGCCTGTGGG 226
|||||
DB 876 TGGGTGGAGAGCAGCGCAACTCGGGCGTGGCGGTGGCGAGGGAGCGCTGCGCCTGTGGG 817

QY 227 ATTCACAGTATGCGAGTGCCTGCTCAGCCACCGCTGGGCACCTCAGCGGGCGGCACCT 286
|||||
DB 816 ATTCACAGTATGCGAGTGCCTGCTCAGCCACCGCTGGGCACCTCAGCGGGCGGCACCT 757

QY 287 GCTTTGAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATGTCAGTTTGGCCAGC 346
|||||
DB 756 GCTTTGAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATGTCAGTTTGGCCAGC 697

QY 347 TGACTTCACATGCATCCTCTTGGAGCCTGCGAGCGCTACTACACCGGTTACTTCGTATCGA 406
|||||
DB 696 TRACTTCATGCCATCCTCTTGGAGCCTGCGAGCGCTACTACACCGGTTACTTCGTATCGA 637

QY 407 ATATCTATCTGAGCCCTCGCTACCTTGGGGAATTCACCTATGACATTCCTTGGTGAAGC 466
|||||
DB 636 ATATCTATCTGAGCCCTCGCTACCTTGGGGAATTCACCTATGACATTCCTTGGTGAAGC 577

QY 467 TGTCTGACCTGTCACTTACACTAAACACATCCAGCCCATCTGTCACGCGCTCCACAT 526
|||||
DB 576 TGTCTGACCTGTCACTTACACTAAACACATCCAGCCCATCTGTCACGCGCTCCACAT 517

QY 527 TTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGCTGGGGGTATCATCAAGAGGATG 586
|||||
DB 516 TTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGCTGGGGGTATCATCAAGAGGATG 457

QY 587 AGGCACCTGCCATCTCCACACCCCTCCAGGAAGTTTCAGGTCCGCATCATATAACACTTA 646
|||||
DB 456 AGGCACCTGCCATCTCCACACCCCTCCAGGAAGTTTCAGGTCCGCATCATATAACACTTA 397

QY 647 TGTGCAACCACTCTCTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGGTTT 706
|||||
DB 396 TGTGCAACCACTCTCTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGGTTT 337

QY 707 GTCTGGCAATGCCAAGCGGGAAGATGCTGCTGCTGGTGAAGTACAGTGGACCTTGG 766
|||||
DB 336 GTCTGGCAATGCCAAGCGGGAAGATGCTGCTGCTGGTGAAGTACAGTGGACCTTGG 277

QY 767 CTTGTAACAGATGAGTGTGTATCATGATTTGAGTTCAGTCTGGAGTGGGGCTGTG 826
|||||
DB 276 CTTGTAACAGATGAGTGTGTATCATGATTTGAGTTCAGTCTGGAGTGGGGCTGTG 217

QY 827 GTCCGCCAATCGGCCCGGTGTCTACCAATATCAGCCACCACTTTGAGTGGATCCAGA 886
|||||
DB 216 GTCCGCCAATCGGCCCGGTGTCTACCAATATCAGCCACCACTTTGAGTGGATCCAGA 157

QY 887 AGCTGATGCCAGAGTGGCATGTCCAGCAGACCCCTCTCGGCACCTACTCTTTTCC 946
|||||
DB 156 AGCTGATGCCAGAGTGGCATGTCCAGCAGACCCCTCTCGGCAGCTACTCTTTTCC 97

QY 947 CTCCTCTCTGGGCTCTCCACCTCTCTGGGCGGCTGTGAGCCTACCTGAGCCCATCGACC 1006
|||||
```

```
DB 96 CTCTTCHCTGGGCTCTCCCMCTCTGGGGCGGCTGTGAGCTACCTGAGCCCATGCGAGCC 37
QY 1007 TGGGGCCACTGCGCAAGTCAGGCC 1029
|||||
DB 36 TGGGGCCACTGCGCAAGTCAGGCC 14

RESULT 2
AL555870 876 bp mRNA linear EST 16-FEB-2001
LOCUS AL555870 LTI_NFL006.PL2 Homo sapiens cDNA clone CSODK002YM07 5
DEFINITION prime, mRNA sequence.
ACCESSION AL555870
VERSION AL555870.1 GI:12898012
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 876)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODK002YM07"
/tissue_type="placenta"
/notes="vector: pCMVSPORT 6; Site.1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched; double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 171 a 251 c 266 g 188 t
ORIGIN
```

Query Match 79.8%; Score 862.4; DB 9; Length 876;  
Best Local Similarity 99.8%; Pred. No. 3.8e-212;  
Matches 874; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 34 GAGGCCATGGCGCGCGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 93
|||||
DB 1 GAGGCCATGGCGCGCGCGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60

QY 94 AGGAACCCGAGTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 152
|||||
DB 61 AGGAACCCGAGTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120

QY 153 CACGTCCGCATCTGCTGGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 212
|||||
DB 121 CACGTCCGCATCTGCTGGGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180

QY 213 CTTGGCGCTGTGGGATTTCCACGATATGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 272
|||||
DB 181 CTTGGCGCTGTGGGATTTCCACGATATGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

QY 273 CACGCGCGGCACTGCTTTGAACCTATAGTACCTTAGTATCCCTCGGGGTGGATGT 332
|||||
DB 241 CACGCGCGGCACTGCTTTGAACCTATAGTACCTTAGTATCCCTCGGGGTGGATGT 300

QY 333 CCAGTTTGGCCAGCTGACTTCCATGCCATCTCTTCGAGCCTGCGAGGCTACTACACCGG 392
|||||
```

```
Db 301 CCAGTTTGGCAGCTGACTTCCATGCCATCCTCTGGAGCCTGCGAGGCTACTACACCG 360
Qy 393 TTACTTGTATGATATCTATCTAGCCCTCGCTACTCTGGGGAATTCACCCCTATGACAT 452
Db 361 TTACTTGTATGATATCTATCTAGCCCTCGCTACTCTGGGGAATTCACCCCTATGACAT 420
Qy 453 TGCCTTGGTGAAGCTGTCTGACCTGTCACTACATCACTAAACACATCCAGCCCATCTGTCT 512
Db 421 TGCCTTGGTGAAGCTGTCTGACCTGTCACTACATCACTAAACACATCCAGCCCATCTGTCT 480
Qy 513 CCAGGCTCCACATTTGAGTTTGAAGCCGACAGACTGCTGGTGACTGCTGGGGTA 572
Db 481 CCAGGCTCCACATTTGAGTTTGAAGCCGACAGACTGCTGGTGACTGCTGGGGTA 540
Qy 573 CATCAAGAGATGAGGCACTGCCATCTCCACACACCTCCAGGAAGTTCAAGTCGCCAT 632
Db 541 CATCAAGAGATGAGGCACTGCCATCTCCACACACCTCCAGGAAGTTCAAGTCGCCAT 600
Qy 633 CATAAACAACCTCTATGTGCAACACCTCTCTCTCAAGTACAGTTTCCGCAAGGACATCTT 692
Db 601 CATAAACAACCTCTATGTGCAACACCTCTCTCTCAAGTACAGTTTCCGCAAGGACATCTT 660
Qy 693 TGGAGACATGTTTGTCTGTCATATGCCAAGCGGGAGGATGCTGCTTCGCTGACTC 752
Db 661 TGGAGACATGTTTGTCTGTCATATGCCAAGCGGGAGGATGCTGCTTCGCTGACTC 720
Qy 753 AGGTGGACCTTGGCTGTATCAAGAATGGACTGTGGTATCAGATTTGAGTCGTGAGCTG 812
Db 721 AGGTGGACCTTGGCTGTATCAAGAATGGACTGTGGTATCAGATTTGAGTCGTGAGCTG 780
Qy 813 GGGAGTGGGCTGTGGTGGGCCAATCGGCCGGGTGTCTTACACCAATATACGCCACCAT 872
Db 781 GGGAGTGGGCTGTGGTGGGCCAATCGGCCGGGTGTCTTACACCAATATACGCCACCAT 840
Qy 873 TGAGTGGATCAGAAAGCTGATGGCCCGAGTGGCAT 908
Db 841 TGAGTGGATCAGAAAGCTGATGGCCCGAGTGGCAT 876
```

## RESULT 3

```
BI334059 602997320F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139509 5',
LOCUS mRNA sequence.
```

```
ACCESSION BI334059
```

```
VERSION BI334059.1 GI:15018703
```

```
KEYWORDS EST.
```

```
SOURCE human.
```

```
ORGANISM Homo sapiens
```

```
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
```

```
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
JOURNAL Unpublished (1999)
```

```
COMMENT Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-r@mail.nih.gov
```

```
Tissue Procurement: ATCC
```

```
cDNA Library Preparation: Life Technologies, Inc.
```

```
cDNA Library Arrayed by: Incyte Genomics, Inc.
```

```
DNA Sequencing by: Incyte Genomics, Inc.
```

```
Clone Distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: LUAM1343 row: h column: 06
```

```
High quality sequence stop: 826.
```

## FEATURES

```
Location/Qualifiers
```

```
1..871
```

```
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
```

```
/clone_image="IMAGE:5139509"
```

```
/clone_lib="NIH_MGC_12"
```

```
/tissue_type="cervical carcinoma cell line"
```

```
BASE COUNT 185 a 258 c 224 g 204 t
ORIGIN
Query Match 71.9%; Score 777.2; DB 13; Length 871;
Best Local Similarity 97.9%; Pred. No. 4e-190;
Matches 842; Conservative 0; Mismatches 8; Indels 10; Gaps 5;
```

```
Qy 209 GGAGCCTCGCCTGTGGGATTCACACGATGCGGAGTGGAGCTGCTAGCCACCGCTGG 268
```

```
Db 1 GGAGCCTCGCCTGTGGGATTCACACGATGCGGAGTGGAGCTGCTAGCCACCGCTGG 60
```

```
Qy 269 CACTCACGGCGGCGGACTGTCTTGAACCTATAGTGACCTTAGTGATCCCTCCGGGTGA 328
```

```
Db 61 CACTCACGGCGGCGGACTGTCTTGAAC-----TGACCTTAGTGATCCCTCCGGGTGA 114
```

```
Qy 329 TGGTCCAGTTTGGCCAGCTGACTTCCATGCTCATCTCTGGAGCCTGCGAGCCTACTACA 388
```

```
Db 115 TGGTCCAGTTTGGCCAGCTGACTTCCATGCTCATCTCTGGAGCCTGCGAGCCTACTACA 174
```

```
Qy 389 CCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACTCGGGGAATTCACCCATG 448
```

```
Db 175 CCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACTCGGGGAATTCACCCATG 234
```

```
Qy 449 ACATTGCCCTTGGTGAAGCTGTCTGCACTGCTACCTACACTAAACACATCCAGCCCATCT 508
```

```
Db 235 ACATTGCCCTTGGTGAAGCTGTCTGCACTGCTACCTACACTAAACACATCCAGCCCATCT 294
```

```
Qy 509 GTCTCCAGGCTCCACATTTGAGTTTGAACCGGAGACTGCTGGGTGACTGGCTGGG 568
```

```
Db 295 GTCTCCAGGCTCCACATTTGAGTTTGAACCGGAGACTGCTGGGTGACTGGCTGGG 354
```

```
Qy 569 GGTACATCAAGAGAGATGAGGACTGCGCATCTCCCCACACCTCCAGGAAGTTTCAGGTCG 628
```

```
Db 355 GGTACATCAAGAGAGATGAGGACTGCGCATCTCCCCACACCTCCAGGAAGTTTCAGGTCG 414
```

```
Qy 629 CCATCATAAACAACCTCTATGTGCAACACCTCTTCTCTCAAGTACAGTTTCGCAAGGACA 688
```

```
Db 415 CCATCATAAACAACCTCTATGTGCAACACCTCTTCTCTCAAGTACAGTTTCGCAAGGACA 474
```

```
Qy 689 TCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGGCGGAAGGATGCTGCTTCGGGTG 748
```

```
Db 475 TCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGGCGGAAGGATGCTGCTTCGGGTG 534
```

```
Qy 749 ACTCAGTGGACCTTGGCCTGTAAACAAGATGGACTGTGGTATCAGATTGG-AGTCGTG 807
```

```
Db 535 ACTCAGTGGACCTTGGCCTGTAAACAAGATGGACTGTGGTATCAGATTGGCAGTCGTG 594
```

```
Qy 808 AGCTGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACCAATATCAGCCAC 867
```

```
Db 595 AGCTGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACCAATATCAGCCAC 654
```

```
Qy 868 CACTTTGAGTGGATCCAGAAGCTGATGCCAGAGTGGCATGTCCACGACAGACCCCTCC 927
```

```
Db 655 CACTTTGAGTGGATCCAGAAGCTGATGCCAGAGTGGCATGTCCACGACAGACCCCTCC 714
```

```
Qy 928 TGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACCTCTGGGCGCGGTCTGAGCC 987
```

```
Db 715 TGGCC-GTACTCTTTTCCCTCTCTCTGGGCTCTCCACCTCTGGGCGCGGTCTGAGCC 773
```

```
Qy 988 TACCTGAGCCCATGAGCCTGGGCGCACTGCCAAGTACAGCCCTGGTCTCTTCTGTCTT 1047
```

```
Db 774 TAACTGAGCCCATGAGCCTGGGCGCACTGCCAAGTACAGCCCTGGTCTCTTCTGTCTT 832
```

```
Qy 1048 GTTTGGTAAATAACACATTC 1067
```

```
Db 833 G-TAGGTAATAAAACATTC 851
```

RESULT 4	BI259237	777 bp	mRNA	linear	EST 17-JUL-2001
LOCUS	602970095F1 NIH_MGC_12	Homo sapiens	CDNA clone	IMAGE:5109695 5'	
DEFINITION	mRNA sequence.				
ACCESSION	BI259237				
VERSION	BI259237.1	GI:14816376			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1265 row: m column: 24 High quality sequence stop: 777. Location/Qualifiers				
FEATURES	1..777				
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5109695" /clone_lib="NIH_MGC_12" /tissue_type="cervical carcinoma cell line" /lab_host="DH10B" /note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."				
BASE COUNT	159 a 224 c 224 g 170 t				
ORIGIN					
Query Match	70.3%;	Score 760.4;	DB 13;	Length 777;	
Best Local Similarity	99.7%;	Pred. No. 8.4e-186;			
Matches 772;	Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1;
QY	91	CTCAGGAAGCCGGAGTTCGAGGAGGCGGCCCTTATCAGGACCATGCGCGGATC	150		
Db	1	CTCAGGAAGCCGGAGTTCGAGGAGGCGGCCCTTATCAGGACCATGCGCGGATC	60		
QY	151	ATCACGTGCGGCATCGTGGGTGGAGGACGCGCAACTCGGGCGTTGGCGGTGGCAGGGG	210		
Db	61	ATCACGTGCGGCATCGTGGGTGGAGGACGCGCAACTCGGGCGTTGGCGGTGGCAGGGG	120		
QY	211	AGCGTCGCGCTCTGGGATTCGCCACGTATGCGAGTGGACCTGCTCAGCCACCGCTGGGCA	270		
Db	121	AGCGTCGCGCTCTGGGATTCGCCACGTATGCGAGTGGACCTGCTCAGCCACCGCTGGGCA	180		
QY	271	CTCACGGCGCGCACTGCTTTGAACACCTATAGTGACCTTAGTGATCCCTCGGGGTGGATG	330		
Db	181	CTCACGGCGCGCACTGCTTTGAACACCTATAGTGACCTTAGTGATCCCTCGGGGTGGATG	240		
QY	331	GTCCAGTTGGCGCAGCTGACTTCCATGCCATCCCTTCGGAGCCTGACAGGCTACTACACC	390		
Db	241	GTCCAGTTGGCGCAGCTGACTTCCATGCCATCCCTTCGGAGCCTGACAGGCTACTACACC	300		
QY	391	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCATGAC	450		
Db	301	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCATGAC	360		
QY	451	ATTGCGTTGGTGAAGCTGCTGCGACCTGTCACCTACACTAAACACATCCAGCCATCTGT	510		
Db	361	ATTGCGTTGGTGAAGCTGCTGCGACCTGTCACCTACACTAAACACATCCAGCCATCTGT	420		

```
Query Match 67.8%; Score 733.2; DB 13; Length 878;
Best Local Similarity 94.4%; Pred. No. 9.4e-179;
Matches 859; Conservative 0; Mismatches 13; Indels 38; Gaps 8;

QY 62 TCGTGGCGCTGCTGCTGGCTGGCTGGCTGACTCAGGAACCGGAGTCGAGAGCGGCGC 121
DB 1 TCGTGGCGCTGCTGCTGGCTGGCTGGCTGACTCAGGAACCGGAGTCGAGAGCGGCGC 60
QY 122 CCTTATCAGGACCATCGCGCGGAGGTCATCAGCTCCGCGATCGTGGTGGAGAGGACG 181
DB 61 CGTTATCAGGACCATCGCGCGGAGGTCATCAGCTCCGCGATCGTGGTGGAGAGGACG 120
QY 182 CCGAATCGGCGGTTGGCGGTGGAGAGGAGCGCTGCGCGCTGCGGATTCACAGTATCGG 241
DB 121 CCGAATCGGCGG-TGGCGGTGGAGGAGGAGCGCTGCGCGTGGGATTCACAGTATCGG 179
QY 242 GAGTGAGCGCTGCTCAGCCACCGCTGGGCACATCAGCGGCGGCGACTGCTTGAACCTATA 301
DB 180 GAGTGAGCGCTGCTCAGCCACCGCTGGGCACATCAGCGGCGGCGACTGCTTGAAC 234
QY 302 GTGACCTTAGTATCCCTCCGCGGTGGATGTCAGTGGTGGCGAGCTGACTTCCATGCCAT 361
DB 235 -TGACCTTAGTATCCCTCCGCGGTGGATGTCAGTGGTGGCGAGCTGACTTCCATGCCAT 293
QY 362 CTTCTGGAGCGCTGAGGCGCTACTACACCCGTTACTTCGTATCGAATATCATCTGAGCC 421
DB 294 CTTCTGGAGCGCTGAGGCGCTACTACACCCGTTACTTCGTATCGAATATCATCTGAGCC 353
QY 422 CTCGCTACCTGGGNAATCACCCATGACATTCGCTGGTGAAGCTGCTGCACCTGTCA 481
DB 354 CTCGCTACCTGGGNAATCACCCATGACATTCGCTGGTGAAGCTGCTGCACCTGTCA 413
QY 482 CCTACATTAACACATCCAGCCCATCTGCTCCAGGCGCTCCACATTTGAGTTGAGAACC 541
DB 414 CCTACATTAACACATCCAGCCCATCTGCTCCAGGCGCTCCACATTTGAGTTGAGAACC 473
QY 542 GGACAGACTGCTGGTGAAGTGGTGGGTATACATCAAGAGAGATGAGGCACTGCCATCTC 601
DB 474 GGACAGACTGCTGGTGAAGTGGTGGGTATACATCAAGAGAGATGAG 520
QY 602 CCCACACCTCCAGGAGTTTCAGTCCGCTCCATTAACCACTCTATGTCGAACCACTCT 661
DB 521 -----GGAAGTTTCAGTCCGCTCCATTAACCACTCTATGTCGAACCACTCT 567
QY 662 TCCTCAAGTACAGTTTCGCGAAGGACATCTTTGGAGACATGTTTGTGCTGGCAATGCC 721
DB 568 TCCTCAAGTACAGTTTCGCGAAGGACATCTTTGGAGACATGTTTGTGCTGGCAATGCC 627
QY 722 AAGCGGGAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
DB 628 AAGCGGGAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 687
QY 780 TGGAGTGTGGTATCAGATTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 838
DB 588 TGGAGTGTGGTATCAGATTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 747
QY 839 GGCCCGGTGTCTACAGCAATATCAGCCACCACTTTGAG-TGGATCCAGAGCTGATGGCC 897
DB 748 GGCCCGGTGTCTACAGCAATATCAGCCACCACTTTGAGTTGATCCAGAGCTGATGGCC 807
QY 898 CAGAGTGCGATGTCAGCCAGCAG-ACCCTCTCTGCGCACTACTCTTTTCCCTCTCTCTG 956
DB 808 CAGAGTGCGATGTCAGCCAGAGCCCTCTCTGCGCACTACTCTTTTCCCTCTCTCTG 867
QY 957 GGCTCTCCCA 966
DB 868 GGCTCTCCCA 877

RESULT 6
AII26185/c 633 bp mRNA linear EST 28-OCT-1998
LOCUS AII26185
DEFINITION qd81h05.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1735929
```

```
3' similar to SW:PSS8_HUMAN Q16651 PROSTASIN PRECURSOR ;, mRNA
sequence.
AII26185
VERSION AII26185.1 GI:3594699
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 633)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: ccapbs-remail.nih.gov
CNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 912 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 455.
FEATURES
Location/Qualifiers
1..633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1735929"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 151 a 164 c 185 g 131 t 2 others
ORIGIN
Query Match 57.8%; Score 624.8; DB 9; Length 633;
Best Local Similarity 99.4%; Pred. No. 8.9e-151;
Matches 626; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 448 GACATTGCTTGGTGAAGCTGCTGCACCTGTGCACCTACACTAAACACATCCAGCCCATC 507
DB 633 GACATTGCTTGGTGAAGCTGCTGCACCTGTGCACCTACACTAAACACATCCAGCCCATC 574
QY 508 TGCTCCAGGCGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGCATCGCTGG 567
DB 573 TGCTCCAGGCGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGCATCGCTGG 514
QY 568 GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCTCCAGAGAGTTTCAGGTC 627
DB 513 GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCTCCAGAGAGTTTCAGGTC 454
QY 628 GCCATATAAACAACCTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCGAAGGAC 687
DB 453 GCCATATAAACAACCTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCGAAGGAC 394
QY 688 ATCTTTGGAGACATGTTTGTGCTGGCAATGCCAAGCGGGGAGAGTGCCTGCTTCGGT 747
DB 393 ATCTTTGGAGACATGTTTGTGCTGGCAATGCCAAGCGGGGAGAGTGCCTGCTTCGGT 334
QY 748 GACTCAGTGGACCCCTTGGCTGTAAAGAAATGAGTGTGTATCAGATTGGAATCGTGTG 807
```

3' similar to SW:PSS8\_HUMAN Q16651 PROSTASIN PRECURSOR ;, mRNA sequence.  
AII26185  
VERSION AII26185.1 GI:3594699  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 633)  
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: ccapbs-remail.nih.gov  
CNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo , Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 912 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 455.

FEATURES  
Location/Qualifiers  
1..633  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1735929"  
/clone\_lib="Soares\_testis\_NHT"  
/sex="male"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories , Inc., and primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 151 a 164 c 185 g 131 t 2 others  
ORIGIN  
Query Match 57.8%; Score 624.8; DB 9; Length 633;..  
Best Local Similarity 99.4%; Pred. No. 8.9e-151;..  
Matches 626; Conservative 0; Mismatches 4; Indels 0; Gaps 0;..  
QY 448 GACATTGCTTGGTGAAGCTGCTGCACCTGTGCACCTACACTAAACACATCCAGCCCATC 507  
DB 633 GACATTGCTTGGTGAAGCTGCTGCACCTGTGCACCTACACTAAACACATCCAGCCCATC 574  
QY 508 TGCTCCAGGCGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGCATCGCTGG 567  
DB 573 TGCTCCAGGCGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGCATCGCTGG 514  
QY 568 GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCTCCAGAGAGTTTCAGGTC 627  
DB 513 GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCTCCAGAGAGTTTCAGGTC 454  
QY 628 GCCATATAAACAACCTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCGAAGGAC 687  
DB 453 GCCATATAAACAACCTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCGAAGGAC 394  
QY 688 ATCTTTGGAGACATGTTTGTGCTGGCAATGCCAAGCGGGGAGAGTGCCTGCTTCGGT 747  
DB 393 ATCTTTGGAGACATGTTTGTGCTGGCAATGCCAAGCGGGGAGAGTGCCTGCTTCGGT 334  
QY 748 GACTCAGTGGACCCCTTGGCTGTAAAGAAATGAGTGTGTATCAGATTGGAATCGTGTG 807

```

Db 333 GACTCAGGTGGACCTTGGCCCTGTAACAAGATGGACTGTGTATCAGATTGGAGTCGTG 274
QY 808 AGCTGGGAGTGGGTGTGGCCCAATCGGCCCGGTGTCTACCAATATATCAGCCAC 867
Db 273 AGCTGGGAGTGGGTGTGGCCCAATCGGCCCGGTGTCTACCAATATATCAGCCAC 214
QY 868 CACTTTGAGTGATCCAGAGCTGATGCCCGCAGAGTGGCATGTCCAGCCAGAGCCCTCC 927
Db 213 CACTTTGAGTGATCCAGAGCTGATGCCCGCAGAGTGGCATGTCCAGCCAGAGCCCTCC 154
QY 928 TGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCCGCGGCGGTCTGAGCC 987
Db 153 TGGCCGCTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCCGCGGCGGTCTGAGCC 94
QY 988 TACTGAGCCATGACGCTGGGGCCACTGCCAAGTCCAGCCCTGTGTCTTCTGTCTT 1047
Db 93 TACTGAGCCATGACGCTGGGGCCACTGCCAAGTCCAGCCCTGTGTCTTCTTCTGTCTT 34
QY 1048 GTTTGGTAATAACACATTCAGTTGATGC 1077
Db 33 GTTTGGTAATAACACATTCAGTTGATGC 4

RESULT 7
BM019026 705 bp mRNA linear EST 30-OCT-2001
LOCUS 603647010F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5428539 5',
DEFINITION mRNA sequence.
ACCESSION BM019026
VERSION BM019026.1 GI:165333380
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 705)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1896 row: c column: 04
High quality sequence start: 7
High quality sequence stop: 675.
Location/Qualifiers
1..705
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5428539"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 139 a 201 c 211 g 154 t
ORIGIN

```

Query Match 52.5%; Score 567.6; DB 13; Length 705;  
 Best Local Similarity 93.0%; Pred. No. 5,6e-136;  
 Matches 649; Conservative 0; Mismatches 19; Indels 30; Gaps 4;

```

QY 50 GCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGCTGAGTCCAGGAAGCCGAGTCGC 109
Db 2 GCTGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGCTGAGTCCAGGAAGCCGAGTCGC 60
QY 110 AGGAGGCGGCGCCCTTATCAGGACCATATCGGCCCGACGGGTATCATCAGCTCGCGCATCGTGG 169
Db 61 A-GAGGCGGCGCCCTTATCAGGACCATATCGGCCCGACGGGTATCATCAGCTCGCGCATCGTGG 119
QY 170 GTGGAGGAGGACGCGAACTCGGGGCTTGGCGGTGGCAGGAGGAGGCTCGGCTGTGGGATT 229
Db 120 GTGGAGGAGGACGCGAACTCGGGGCTTGGCGGTGGCAGGAGGAGGCTCGGCTGTGGGATT 179
QY 230 CCCACGTATCGGAGTGAGCTGTCTAGCCACCGCTGGGCACCTCAGCCGCGGCGGCGACATGCT 289
Db 180 CCCACGTATCGGAGTGAGCTGTCTAGCCACCGCTGGGCACCTCAGCCGCGGCGGCGACATGCT 239
QY 290 TTGAACCTATAGTGACCTTAGTATCCCTCCGGGTGGATGGTCCAGTCTTGGCCAGCTGA 349
Db 240 TTGAACCTATAGTGACCTTAGTATCCCTCCGGGTGGATGGTCCAGTCTTGGCCAGCTGA 299
QY 350 CTTCCATGCCATCCTTCTGGAGCCTGCAGGCTTACTACACCCCTTACTTCTGATTCGATCGAATA 409
Db 300 CTTCCATGCCATCCTTCTGGAGCCTGCAGGCTTACTACACCCCTTACTTCTGATTCGGAATA 359
QY 410 TCTATCTGAGCCCTCGCTACTTGGGAATTCACCCCTATGACATTCCTTGGTGAAGCTGT 469
Db 360 TCTATCTGAGCCCTCGCTACTTGGGAATTCACCCCTATGACATTCCTTGGTGAAGCTGT 419
QY 470 CTGCACCTGTACCTACATTAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTTG 529
Db 420 CTGCACCTGTACCTACATTAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTTG 479
QY 530 AGTTTCAGAACCGACAGACTGCTGGGTGACTGCTGGGGGTACATCAAGAGAGATCAGG 589
Db 480 AGTTTCAGAACCGACAGACTGCTGGGTGACTGCTGGGGGTACATCAAGAGAGATCAGG 538
QY 590 CACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGTCCGCCATCATATAACAACCTATGT 649
Db 539 -----GGAAAGTTTCAGTCCGCCATCATATAACAACCTATGT 573
QY 650 GCACACACCTCTCTCAAGTACAGTTTCGGAAGGACATCTTTGGAGACATGTTTGTG 709
Db 574 GCACACACCTCTCTCTCAAGTACAGTTTCGGAAGGACATCTTTGGAGACATGTTTGTG 633
QY 710 CTGGCAA--TGCCCAAGCGGGAAGATGCTGCTTCG 745
Db 634 CTGGCAAATGCCCAAGCGGGAAGATGCTGCTTCG 671

RESULT 8
BM017444 682 bp mRNA linear EST 30-OCT-2001
LOCUS 603645444F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5426338 5',
DEFINITION mRNA sequence.
ACCESSION BM017444
VERSION BM017444.1 GI:16531798
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1896 row: c column: 04
High quality sequence start: 7
High quality sequence stop: 675.
Location/Qualifiers
1..682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5426338"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 139 a 201 c 211 g 154 t
ORIGIN

```

found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM1890 row: 9 column: 11  
High quality sequence stop: 672.

## FEATURES

source  
1. .682  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5426338"  
/clone\_lib="NIH\_MGC\_98"  
/tissue\_type="astrocytoma grade IV, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/notes="Organ: brain; Vector: pOT87; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dr priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACAGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 135 a 204 c 197 g 146 t  
ORIGIN

Query Match 51.0%; Score 551.4; DB 13; Length 682;  
Best Local Similarity 97.9%; Pred. No. 8.5e-132;  
Matches 569; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
QY 120 GCCCTTATCAGGACCATCGCGCCGACGGTTCATCACGTCGCGCATCGTGGTGAGAGGA 179  
DB 102 GACCATCGGAGACCATCGCGCCGACGGTTCATCACGTCGCGCATCGTGGTGAGAGGA 161  
QY 180 CGCGAACTCGGGGTTGGCC-GTGGCAGGGAGCCCTCGGCTCTGGGATTCACACGAT 238  
DB 162 CGCGAACTCGGGGTTGGCCGTGGCAGGGAGCCCTCGGCTCTGGGATTCACACGAT 221  
QY 239 GCGAGTAGGCTCTCAGCCACCGCTGGGCACTACAGGCGGCGCACTGCTTTGAAACCT 298  
DB 222 GCGAGTAGGCTCTCAGCCACCGCTGGGCACTACAGGCGGCGCACTGCTTTGAAACCT 281  
QY 299 ATAGTGACCTTAGTGATCCCTCGGGTGGATGTCAGTTGGCCAGCTGACTTCCATGC 358  
DB 282 ATAGTGACCTTAGTGATCCCTCGGGTGGATGTCAGTTGGCCAGCTGACTTCCATGC 341  
QY 359 CATCTTCTGAGGCTCGAGGCTTACTACACCCGTTACTTGTATCGAATATCTATCTGA 418  
DB 342 CATCTTCTGAGGCTCGAGGCTTACTACACCCGTTACTTGTATCGAATATCTATCTGA 401  
QY 419 GCCCTCGCTACCTGGGGAATTCACCCATGACATTCGCTTGGTGAAGCTGCTGCACCTG 478  
DB 402 GCCCTCGCTACCTGGGGAATTCACCCATGACATTCGCTTGGTGAAGCTGCTGCACCTG 461  
QY 479 TCACCTACACTAAACACATCCAGCCCATCTCTCCAGGCTCCACATTTGAGTTTGA 538  
DB 462 TCACCTACACTAAACACATCCAGCCCATCTCTCCAGGCTCCACATTTGAGTTTGA 521  
QY 539 ACCGACAGACTGCTGGGTGACTGGCTGGGGTATACAAAGAGGATGAGCACTGCCAT 598  
DB 522 ACCGACAGACTGCTGGGTGACTGGCTGGGGTATACAAAGAGGATGAGCACTGCCAT 581  
QY 599 CTCCGACACCTCCAGGAGTTTCAGTGGCCATATCAACAACTCTATGTGCAACACC 658  
DB 582 CTCCGACACCTCCAGGAGTTTCAGTGGCCATATCAACAACTCTATGTGCAACACC 641  
QY 659 TCTTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGAC 699  
DB 642 TCTTCTCAAGTACAGTTTCCGCAAGGACATCTCGGGAGAC 682

## RESULT 9

BM553333

LOCUS

DEFINITION

BM553333 1219 bp mRNA linear EST 20-FEB-2002  
AGENCOURT\_6558382 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5742958  
5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BM553333  
BM553333.1 GI:18791973  
EST.  
human.  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 1219)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM12761 row: 0 column: 23  
High quality sequence stop: 553.  
Location/Qualifiers  
1. .1219  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5742958"  
/clone\_lib="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dr primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."  
BASE COUNT 225 a 393 c 354 g 246 t 1 others  
ORIGIN

## Query Match

Best Local Similarity 50.9%; Score 549.8; DB 13; Length 1219;  
Best Local Similarity 91.1%; Pred. No. 2.6e-131;  
Matches 606; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

QY 120 GCCCTTATCAGGACCATCGCGCCGACGGTTCATCACGTCGCGCATCGTGGTGAGAGGA 179  
DB 57 GACCATCGGAGACCATCGCGCCGACGGTTCATCACGTCGCGCATCGTGGTGAGAGGA 116  
QY 180 CGCGAACTCGGGGTTGGCCGTGGCAGGGAGCCCTCGGCTCTGGGATTCACACGATG 239  
DB 117 CGCGAACTCGGGGTTGGCCGTGGCAGGGAGCCCTCGGCTCTGGGATTCACACGATG 176  
QY 240 CGGAGTAGGCTCTCAGCCACCGCTGGGCACTCACGCGCGCACTGCTTTGAAACCTA 299  
DB 177 CGGAGTAGGCTCTCAGCCACCGCTGGGCACTCACGCGCGCACTGCTTTGAAACCTA 236  
QY 300 TAGTGACCTTAGTGATCCCTCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCC 359  
DB 237 TAGTGACCTTAGTGATCCCTCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCC 296  
QY 360 ATCTTCTGGAGCTCGAGGCTTACTACACCCGTTACTTGTATCGAATATCTATCTGAG 419  
DB 297 ATCTTCTGGAGCTCGAGGCTTACTACACCCGTTACTTGTATCGAATATCTATCTGAG 356  
QY 420 CCCTTCGTACTCTGGGGAATTCACCCATATGACATTCGCTTGGTGAAGCTGCTGCACCTGT 479  
DB 357 CCCTTCGTACTCTGGGGAATTCACCCATATGACATTCGCTTGGTGAAGCTGCTGCACCTGT 416  
QY 480 CACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTTGAGTTGAGAA 539  
DB 417 CACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTTGAGTTGAGAA 476



QY	48	CGCGGGCGCGTCTGCTGGCGCTGCTGCTGCTGGCGCTGACATCAGGAAGCC--GGAGT	106
Db	2	GCTGTGGCGCTGCTGCTGGCGCTGCTGCGGCTCGGCTGACATCAGGAAGCCGTGGAGT	61
QY	107	CGCAGAGCGCGCGCCCTTATCAGGACCATGCGGCGCAGGGGTATCAGTGCAGCATCG	166
Db	62	CGCAGAGCGCGCGCTTATCAGGA--CATGCGGCGCAGGGGTATCAGTGCAGCATCG	120
QY	167	TGGGTGGAGAGAGACCGCGAATCTGGCGGCTTGGCCGTGGCAGGGAGCCCTGCGCCTGTGGG	226
Db	121	TGGGTGGAGAGAGACCGCGAATCTGGCGGCTTGGCCGTGGCAGGGAGCCCTGCGCCTGTGGG	180
QY	227	ATTCCACGATGCGGAGTGAGCCCTGATCAGCCACCGCTGGCAGCTAC--GCGGCGGCAC	285
Db	191	ATTCCACGATGCGGAGTGAGCCCTGATCAGCCACCGCTGGCAGCTACAGAGCGCGGCAC	240
QY	286	-TGCTTTGAAACCTATAGTAGCTTACCTTGTGATCCCTCGGGGTGGATGCTTGGGCA	344
Db	241	ATGCTTTGAAACCTATAGTAGCTTACCTTGTGATCCCTCGGGGTGGATGCTTGGGCA	300
QY	345	GCTGAC--TTCCATGCCATCTCTGGAGCCCTGAGCCCTACTACACCCGT--TACTTCTGA	402
Db	301	GCTGACATTCATGCCATCTCTGGAGCCCTGAGCCCTACTACACCCGTATACTTCGTA	360
QY	403	TCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCTATGACATTCCTTGGTG	462
Db	361	TCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCTATGACATTCCTTGGTG	420
QY	463	AAGCTGTCTGACCTGTACCTTACATTAACACATCAGGCCATCTGTCTCCAGGCCCTCC	522
Db	421	AAGCTGTCTGACCTGTACCTTACATTAACACATCAGGCCATCTGTCTCCAGGCCCTCC	480
QY	523	ACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGGTGAGGTGAGGTGAGGTGAG	582
Db	481	ACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGGTGAGGTGAGGTGAGGTGAG	540
QY	583	GATGAGGACTGCGCATCTCCCCACACCCCTCAGGAAGTTTCAAGTGCATCATATAA--CAA	641
Db	541	GATGAGG-----GAAAGTACAGGTGCGCATCATATAAACCACAA	575
QY	642	CTCTATGTGCAACC--ACCTCTTCTCAAGTACACTTT--CCGCAAGGACATCTTTGGAGAC	699
Db	576	CTCTATGTGCAACCACCTCTTCTCAAGTACACTTTCCCGCAAGGACATCTATGGAGAC	635
QY	700	ATGTTTTGTGCTGCAATGCCAAGCGGGAAGATGCTGCTTGGGTGACTCAGGTGGA	759
Db	636	ATGTTTTGTGCTGCAATGCCAAGCGGGAAGATGCTGCTTGGGTGACTCAGGTGGA	695
QY	760	CCCTTGGCTGTAAACAGATGGAATCTGTG	788
Db	696	CACCACTCACAGACCGAGGAAGCATCTCTG	724
RESULT	11		
LOCUS	AA161187/c		
DEFINITION	zo5906.s1 Stratiogene pancreas (#937208) Homo sapiens cDNA clone IMAGE:591155 3' similar to TR:E218488 E218488 TRYPTASE ;, mRNA sequence.	591 bp	linear EST 16-DEC-1996
ACCESSION	AA161187		
VERSION	AA161187.1	GI:1735441	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 591)		
	Chissoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W., Hawkins, B., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, M., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.		



```

|||||
Db 302 CCAGTCACTTCCATCCCTTCTCGAGCCTCGAGCCTACTACACCCGTTACTTCGT 361
QY 402 ATCAATATCTATCTGAGCCCTCGCTACCTGGGAATTCACCCATGACATTCGCTGGT 461
Db 362 ATCAATATCTATCTGAGCCCTCGCTACCTGGGAATTCACCCATGACATTCGCTGGT 421
QY 462 GAAGCTCTGACCTCTGACCTACCTACCTAAACACATCCAGCCCATCTGCTCCAGGCCCTC 521
Db 422 GAAGCTCTGACCTCTGACCTACCTAAACACATCCAGCCCATCTGCTCCAGGCCCTC 481
QY 522 CACATTTGATTTGAGAACCGGACAGACTGCTGGGTGAGCTGCTGGGGTACATC 576
Db 482 CACATTTGATTTAGCAACCGGACAGACTGCTGGGTGAGCTGCTGGGGTCTCCTC 536

RESULT 13
BI257649
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11259 row: m column: 17
High quality sequence stop: 753.
Location/Qualifiers
1..825
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5107384"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT 162 a 245 c 237 g 181 t
ORIGIN
Query Match 46.4%; Score 501.6; DB 13; Length 825;
Best Local Similarity 95.4%; Pred. No. 6.7e-119;
Matches 605; Conservative 0; Mismatches 14; Indels 15; Gaps 8;
QY 120 GCCCTTATCAGCACCATCGGCGGAGGTGTCATCAGTCGCGCATC-TGTGGGTGAGAGG 178
Db 119 GACCATCGAGGACCATCGCGCGAGGTGTCATCAGTCGCGCATC-TGTGGGTGAGAGG 178
QY 179 AC-GCCGAACCTGGCGGTTGGCGGAGGAGCCCTGCGCTGTGGATTCACACCTA 237
Db 179 ACTCGCGAACCTGGCGGTTGGCGGAGGAGGAGCCCTGCGCTGTGGATTCACACCTA 237
QY 238 TCCGAGTGTAGCTCTCAGCAACCGCTGGGCACTCAGCGGGGCGCACTGCTTTGAAC 297
Db 238 TCCGAGTGTAGCTCTCAGCAACCGCTGGGCACTCAGCGGGGCGCACTGCTTTGAAC 297

```

```

QY 298 TATAGTCACTTAGTGTATCCCTCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATG 357
Db 297 -----TGACCTTAGTGTATCCCTCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATG 351
QY 358 CCATCTTCTTGGAGCCTGCGAGGCTTACTACACCCGTTACTTCTGATCGAATATCTATCTG 417
Db 352 CCATCTTCTTGGAGCCTGCGAGGCTTACTACACCCGTTACTTCTGATCGAATATCTATCTG 411
QY 418 AGCCCTCGCTACCTGCGGGAATTCACCTATGACATTCCTTGGTGAAGCTGCTGCACCT 477
Db 412 AGCCCTCGCTACCTGCGGGAATTCACCTATGACATTCCTTGGTGAAGCTGCTGCACCT 471
QY 478 GTCACTTACACTAAACACATCCAGCCCATCTCTCCAGGCCCTCCACATTTGAGTTGAG 537
Db 472 GTCACTTACACTAAACACATCCAGCCCATCTCTCCAGGCCCTCCACATTTGAGTTGAG 531
QY 538 AACCGGACAGACTGCTGGGTGACTGGC-TGGGGGTACATCAAGAGGAT-GAGGCACCTG 595
Db 532 AACCGGACAGACTGCTGGGTGACTGGC-TGGGGGTACATCAAGAGGATGAGGCACCTG 591
QY 596 CATCTCCCCACACCTCCAGGAAGTTCAGGTCCGATCAATCAACAACTCTATGTCAACC 655
Db 592 CATCTCCCCACACCTCCAGGAAGTTCAGGTCCGATCAATCAACAACTCTATGTCAACC 651
QY 656 ACCTTCTCTCAAGTACAGTTT-CCGCAAGGACATCTTTGGAG--ACATGGTTTGTGCT 711
Db 652 ACCTTCTCTCAAGTACAGTTTTCGCGAAGGACATCTTTGGAGGACATGGTTTGTGCT 711
QY 712 GGCATGCCCCAAGCGGGAAGGATGCCCTGCTTCG 745
Db 712 GGCATGCCCCAAGCGGGAAGGATGCCCTGCTTCG 745

RESULT 14
AK006271
LOCUS
DEFINITION
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700023E12:protease, serine, 21, full insert
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
HTC; CAP trapper
Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1700023E12.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ORGANISM
REFERENCE
1
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed sequencer

```



QY	696	AGACATGTTGTTGTCTGGCAATGCCAAGCGGAGGATGCTGTTGGTGACTCAGG	755
Db	682	AGACATGTTTGGCGTGGCACTCCTGAGGTGCAAGGATGCTGTTTGGTGACTCGGG	741
QY	756	TGGACCCCTTGCCCTGTAAACAAGATGGACTGTGTATCATAGATTGGAGTCGTGAGCTGGGG	815
Db	742	AGGACCCCTTGCCCTGCGACACAGGATACGGTGTGGTATCAGTTGGAGTTGGAGCTGGGG	801
QY	816	AGTGGGCTGTGGTGGGCCCAATCGGCCGGTGTCTACACAAATATCAGCCACCATTGGA	875
Db	802	AATAGGCTGTGGTGGCCAAATTCGCCCTGGAGTCTATACCAACATCAGTCATCACTACAA	861
QY	876	GTGGATCCAGAAGCTGATGCGCCAGAGTGGCATGTCCAGCAGCAGACCCCTCTGGCCACT	935
Db	862	CTGGATCCAGTCAACCATGATCGGCAATGGGCTGCTCAGGCCCTGACCCAGTCCCCCTGCT	921
QY	936	ACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGGCGGTCTCTGAGCTACCTGAG	995
Db	922	ACTGTTTCTACCTCTGGGCTGAGCTTCCTCTTGTCTGAGGCCCTGCCTGAGCCCAACAGTG	981
QY	996	CCCATGACGCTTGGGGCCACTGCCAAGTCAGGCCCTGTTCTCTCTGCTCTGTTTGGTA	1055
Db	982	TACGTCACACCTGTG-----AGGTACGGGTGTCTCTTTTGTATCTTGTCTGCTA	1032
QY	1056	ATAAACACATTT	1066
Db	1033	ATAAACCTGTT	1043

RESULT 15  
 LOCUS BM018436  
 DEFINITION 603646140F1 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5428064 5',  
 mRNA sequence.  
 ACCESSION BM018436  
 VERSION  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I. M. A. G. E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I. M. A. G. E. Consortium/LLNL at:  
<http://image.llnl.gov/>

```

FEATURES
source
High quality sequence stop: 622.
Location/Qualifiers
1..624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5428064"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

BASE COUNT  
ORIGIN

123 a	183 c	185 g	133 t
-------	-------	-------	-------

Note: this is a NIH\_MGC Library.

Query Match	44.7%	Score 483.6;	DB 13;	Length 624;
Best Local Similarity	97.9%	Pred. NO. 2.8e-114;		
Matches '511; Conservative	0;	Mismatches	9;	

Qy	120	GCCTTATCAGGACCATCGCGCCGACGGTCTATCACGTGCGGCATCTGGTGGAGAGGA	179
Db	100	GACCATCGGAGGACCATCGCGCGACGGTCTATCACGTGCGGCATCTGGTGGAGAGGA	159
Qy	180	CGCCGAACCTCGGGCGTTGGCGCTGGCGAGGAGCCTCGCCTGTGGGATCCCACTATG	239
Db	160	CGCGAACTCGGGCGTTGGCGCTGGCGAGGAGCCTCGCCTGTGGATTCCCACCGTATG	219
Qy	240	CGGAGTGAGCGCTGCTCAGCCACCGCTGGCGACTACGGGGGCGCACTGCTTTGAAACCTA	299
Db	220	CGAGGTGAGCCTGCTCAGCCACCGCTGGGCACTACGGGGGCGCACTGCTTTGAAACCTA	279
Qy	300	TAGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTCACTTCGATGCC	359
Db	280	TAGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTCACTTCGATGCC	339
Qy	360	ATCCTTCTCGAGGCTCGAGGCTACTACACCGGTTACTTCGTTATCGAATATCTATCTGAG	419
Db	340	ATCCTTCTCGAGGCTCGAGGCTACTACACCGGTTACTTCGTTATCGAATATCTATCTGAG	399
Qy	420	CCCTCGCTACCTCGGGGAATTCACCCATAGCAATTCGCTTGGTGAAGCTGTCTGCACCTGT	479
Db	400	CCCTCGCTACCTCGGGGAATTCACCCATAGCAATTCGCTTGGTGAAGCTGTCTGCACCTGT	459
Qy	480	CACCTACACTAAACATCCAGCCCATCTGTCTCAGGCGTCCACATTTGAGTTTGAGAA	539
Db	460	CACCTACACTAAACATCCAGCCCATCTGTCTCAGGCGTCCACATTTGAGTTTGAGAA	518
Qy	540	CGGACAGACTGCTGGGTGACTGCTGGGGTACATCAAGAGGATGAGGCACCTGCCATC	599
Db	519	CGGACAGACTGCTGGGTGACTGCTGGGGTACATCAAGAGGATGAGGCACCTGCCATC	578
Qy	600	TCCCCACACCCCTCCAGGAAGTTCAAGTTCGGCTGCCATCATAAACAA	641
Db	579	TCCCCACA-CCTCAGGAAGTTCAAGTTCGGCTGCCATCATAAACAA	619

Search completed: May 10, 2003, 04:57:12  
Job time : 2084 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 02:50:58 ; Search time 351 Seconds  
(without alignments)  
6935.642 Million cell updates/sec

Title: US-09-787-844-1

Perfect score: 1081

Sequence: 1 agaggcagagggggcgctca.....acattccagttgatgcctgc 1081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

```

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	100.0	1081	20	AA87151 Human protease HUP
2	1081	100.0	1081	21	AAA08505 Human cancer-spec1
3	1068.8	98.9	1087	21	AAAC2015 Hydrophobic domain
4	1063.2	98.4	1131	22	AAS01698 Human extracellular
5	1063.2	98.4	1131	22	AAS01706 TADG-16 antisense
6	1055.8	97.7	1100	20	AA52259 Protein PRO303 cDN
7	1055.8	97.7	1100	22	AA72417 Human PRO303 cDNA
8	1054.8	97.6	1077	24	AA27418 Human eosinophil s
9	1052.4	97.4	1079	21	AA245673 Nucleotide sequenc

10	1049.6	97.1	1100	19	AAV59119	Nucleotide sequenc
11	1048.8	97.0	1117	19	AAV59118	Nucleotide sequenc
12	1045.8	96.7	1082	20	AAV5337	Nucleic acid encod
13	1045.8	96.7	1082	24	AA27417	Human eosinophil s
14	1037.8	96.0	1082	22	AA52648	Human secreted pro
15	938.8	86.8	942	20	AAV5336	Nucleic acid encod
16	938.8	86.8	942	21	AAAG2005	Hydrophobic domain
17	752.6	69.6	1103	22	AAV5271	Nucleotide sequenc
18	496.4	45.9	498	22	AA501708	WISH TADG-16 catal
19	494.8	45.8	498	22	AA501707	WISH TADG-16 catal
20	493.4	45.6	529	24	AA561822	Lung small cell ca
21	479	44.3	959	19	AAV59132	Nucleotide sequenc
22	475	43.9	918	22	AAI67198	Nucleotide sequenc
23	465.4	43.1	575	22	AA524506	Human ovarian PCR-
24	465.4	43.1	575	22	AAH83114	Human ovarian tumo
25	422.2	39.1	888	24	ABK31769	DNA encoding novel
26	413.4	38.2	1018	24	AA27415	Human eosinophil s
27	334.2	30.9	3866	19	AAV59133	Genomic sequence o
28	329.8	30.5	837	24	AA27419	Human eosinophil s
29	318.2	29.4	373	24	ABL80716	Human ovarian canc
30	165	15.3	1110	22	AAF76994	Human protease T c
31	165	15.3	1129	22	AA521354	Human cDNA sequenc
32	165	15.3	1151	22	AA208286	Human secreted pro
33	165	15.3	1703	21	AA524473	HTFM clone 337404
34	163.2	15.1	980	19	AAV59136	Nucleotide sequenc
35	161.8	15.0	1212	22	AAI58544	Human polynucleoti
36	161.6	14.9	2457	24	ABK31795	DNA encoding novel
37	157.2	14.5	1157	22	AAI60330	Human polynucleoti
38	155.8	14.4	1130	22	AAF77000	Fusion gene of pro
39	151.8	14.0	1322	21	AA61704	cDNA encoding mous
40	150.6	13.9	2662	24	AA23854	Human protease PRT
41	150.6	13.9	2810	22	AAK94500	Human full-length
42	149	13.8	1958	24	ABK12891	Human protease PRT
43	145.2	13.4	1122	24	AA237036	Mouse tryptase cDN
44	145	13.4	437	24	ABL81643	Human ovarian canc
45	133.4	12.3	393	24	AA27416	Human eosinophil s

#### ALIGNMENTS

```

RESULT 1
ID   AA87151 standard; cDNA; 1081 BP.
XX
AC   AA87151;
XX
DT   27-SEP-1999 (first entry)
XX
DE   Human protease HUPM-3 cDNA.
XX
KW   Protease; human; HUPM-3; cell proliferation; cancer;
XX   immune disorder; inflammation; therapy; ss.
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   CDS 40..984
FT     /tag= a
FT   sig_peptide 40..76
FT     /tag= b
FT   mat_peptide 77..981
FT     /note= "putative signal peptide sequence"
FT     /tag= b
XX
XX   WO9936550-A2.
XX
XX   22-JUL-1999.
XX
XX   12-JAN-1999; 99WO-US00655.
XX
XX   16-JAN-1998; 98US-0008271.
XX

```





PT testicular cancers by measuring expression of a cancer specific gene -  
XX  
PS Claim 6; Page 31; 36pp; English.

The present sequence encodes a human cancer-specific gene (CSG), designated p104. The inventive methods comprise measuring expression of CSG in order to diagnose, stage, image and treat gynecologic and testicular cancers. The levels of CSG are compared to levels found in a normal human control (a change in the measured level of CSG is associated with the presence of the cancer). The early diagnosis of cancers improves the success rate of therapeutic protocols.

Sequence 1081 BP; 202 A; 321 C; 321 G; 237 T; 0 other;

```
Query Match      100.0%; Score 1081; DB 21; Length 1081;
Best Local Similarity 100.0%; Pred. No. 1e-252;
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	AGAGGACAGAGGGGGCGTCAGGCCGGGAGGAGAGGCCATGGCGCGCGCGGGGCGCTGG	60
DB	1	AGAGGCGAGAGGGGGGCTCAGGGCCGGGAGGAGAGGCCATGGCGCGCGCGGGGCGCTGG	60
QY	61	CTGCTGGCGCTGCTGCTGGCTCGGGCTGGACATCAGGAAGCCGAGATCGCAGAGAGCGGG	120
DB	61	CTGCTGGCGCTGCTGCTGGCTCGGGCTGGACTCAGGAAGCCGAGATCGCAGAGAGCGGG	120
QY	121	CCCTTATCAGGACCATCGGSCCAGCAGGGGTATCAGCTCGCGCATCGTGGGTGGAGAGGAC	180
DB	121	CCCTTATCAGGACCATCGGSCCAGCAGGGGTATCAGCTCGCGCATCGTGGGTGGAGAGGAC	180
QY	181	GCCGAACTCGGGGCTTGGCGTGGCAGGGAGGCGCTCGGCTCTGGGATTCGCACATGATGC	240
DB	181	GCCGAACTCGGGGCTTGGCGTGGCAGGGAGGCGCTCGGCTCTGGGATTCGCACATGATGC	240
QY	241	GGAGTGAGCCTGCTCAGCCACCGCTGGGCACATCAGCGCGCGCACATGCTTTGAAACCTAT	300
DB	241	GGAGTGAGCCTGCTCAGCCACCGCTGGGCACATCAGCGCGCGCACATGCTTTGAAACCTAT	300
QY	301	AGTGACCTTATGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA	360
DB	301	AGTGACCTTATGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA	360
QY	361	TCCTTCTGAGGCTCGAGGCGCTACTACCCGGTTACTTCGTATCGAATATCTATCTGAGC	420
DB	361	TCCTTCTGAGGCTCGAGGCGCTACTACCCGGTTACTTCGTATCGAATATCTATCTGAGC	420
QY	421	CCTCGCTACCTGGGGAAATTCACCTATGACATTCGCTTGGTGAAGCTGCTGCACCTGTC	480
DB	421	CCTCGCTACCTGGGGAAATTCACCTATGACATTCGCTTGGTGAAGCTGCTGCACCTGTC	480
QY	481	ACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCGCTCCACATTTGAGTTTGAGAAC	540
DB	481	ACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCGCTCCACATTTGAGTTTGAGAAC	540
QY	541	CGGACAGACTGCTGGGTGACTTGCTGGGGTACATCAAAGAGGATGAGGCACTGCCATCT	600
DB	541	CGGACAGACTGCTGGGTGACTTGCTGGGGTACATCAAAGAGGATGAGGCACTGCCATCT	600
QY	601	CCCCACACCTCCAGGAAGTTCAGGTTCGCCATCATAAACAACTCTATATGCAACCACTTC	660
DB	601	CCCCACACCTCCAGGAAGTTCAGGTTCGCCATCATAAACAACTCTATATGCAACCACTTC	660
QY	661	TTCTCTCAAGTACAGTTTCCGACAGGACATCTTTGGAGACATGGTTTGCTGGCAATGCC	720
DB	661	TTCTCTCAAGTACAGTTTCCGACAGGACATCTTTGGAGACATGGTTTGCTGGCAATGCC	720
QY	721	CAAGGGCGGAAGGATCGCTGCTTCGGTGTACTCAGGTGGACCCCTTGGCCCTGTAAACAAGAT	780
DB	721	CAAGGGCGGAAGGATCGCTGCTTCGGTGTACTCAGGTGGACCCCTTGGCCCTGTAAACAAGAT	780
QY	781	GGACTGTGGTATCAGATTGAGATCGTGAGCTGGGGAGTGGGCTGTGGGCCCAATCGG	840
DB	781	GGACTGTGGTATCAGATTGAGATCGTGAGCTGGGGAGTGGGCTGTGGGCCCAATCGG	840



CC TADG-16 nucleic acid is useful for treating various cancers, including  
CC ovarian, breast, lung, colon and prostate. The TADG-16 nucleic acid,  
CC TADG-16 protein and antibodies specific to TADG-16 are useful for the  
CC diagnosis of cancer. TADG-16 protein or its fragments are useful for  
CC vaccinating an individual against TADG-16. Numerous TADG-16 peptides  
CC (AAU02225-AAU02384) are tested for their binding affinity to the  
CC 8 haplotypes HLA A0201, HLA A0205, HLA A1, HLA A24, HLA B7, HLA B8,  
CC HLA B2702, and HLA B4403.  
XX  
SQ Sequence 1131 BP; 217 A; 332 C; 340 G; 242 T; 0 other;

Query Match 98.4%; Score 1063.2; DB 22; Length 1131;  
Best Local Similarity 99.6%; Pred. No. 2.1e-248;  
Matches 1076; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 AGGAGGAGAGGGGCGCTCAGGCGCGGAGAGAGCCATGGCGCGCGGCGCGCTG 60  
DB 25 AGGAGGAGAGGGGCGCTCAGGCGCGGAGAGAGCCATGGCGCGCGGCGCGCTG 84  
QY 61 CTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 85 CTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 144  
QY 121 CCCTTATCAGACATCCGCGCGGAGCGGCTATCAGCTCGCGCATCTGCTGGTGGAGAGAC 180  
DB 145 CCCTTATCAGACATCCGCGCGGAGCGGCTATCAGCTCGCGCATCTGCTGGTGGAGAGAC 204  
QY 181 GCCGAATCCGGGCTTGGCGCTGCGAGGGAGCTGCGCTGCTGCTGCTGCTGCTGCTG 240  
DB 205 GCCGAATCCGGGCTTGGCGCTGCGAGGGAGCTGCGCTGCTGCTGCTGCTGCTGCTG 264  
QY 241 GGAGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 265 GGAGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 324  
QY 301 AGTGACCTTAGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 325 AGTGACCTTAGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384  
QY 361 TCCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
DB 385 TCCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444  
QY 421 CCTCGCTACCTGGGGAATTCACCTATGACATTCGCTTGGTGAAGCTGCTGCACTGCTC 480  
DB 445 CCTCGCTACCTGGGGAATTCACCTATGACATTCGCTTGGTGAAGCTGCTGCACTGCTC 504  
QY 481 ACCTAGACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTGGAAC 540  
DB 505 ACCTAGACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTGGAAC 564  
QY 541 CGGACAGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCT 600  
DB 565 CGGACAGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCT 624  
QY 601 CCCACACCTCCAGGAAGTTCAGGTCGCCATCATAAACAACCTATGTGCAACACCTC 660  
DB 625 CCCACACCTCCAGGAAGTTCAGGTCGCCATCATAAACAACCTATGTGCAACACCTC 684  
QY 661 TTCTCAAGTACAGTTTCCGAGGACATCTTTGGAGACATGGTTTGGTGGCAATGCC 720  
DB 685 TTCTCAAGTACAGTTTCCGAGGACATCTTTGGAGACATGGTTTGGTGGCAATGCC 744  
QY 721 CAAGGCGGGAAGGATGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCT 780  
DB 745 CAAGGCGGGAAGGATGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCTGCTGGGTGCT 804  
QY 781 GGACTGTGGTATCAGATTGGAGTGGTGGAGTGGGAGTGGGCTGCTGGTGGGCGCCCAATCGG 840  
DB 805 GGACTGTGGTATCAGATTGGAGTGGTGGAGTGGGAGTGGGCTGCTGGTGGGCGCCCAATCGG 864  
QY 841 CCGGGTCTTACACCAATATCAGCCACCATTTGAGTGGATCCAGAGAGCTGATGGCCCGAG 900

DB 865 CCCGGTGTCTACCAATATCAGCCACACACCTTTGAGTGGATCCAGAGCTGATGGCCCGAG 924  
QY 901 AGTGGCATGCTCCAGCCAGAGCCCTCCCTGGCCACTACTCTTTTCCCTCTCTCTGGGCT 960  
DB 925 AGTGGCATGCTCCAGCCAGAGCCCTCCCTGGCCACTACTCTTTTCCCTCTCTCTGGGCT 984  
QY 961 CTCCCACTCTCTGGGCGCGCTGAGCCCTGAGCCCTGAGCCCATGAGCCCTGGGCGC-ACTGCC 1019  
DB 985 CTCCCACTCTCTGGGCGCGCTGAGCCCTGAGCCCTGAGCCCATGAGCCCTGGGCGCACTGCC 1044  
QY 1020 AAGTCAGGCCCTGGTCT 1079  
DB 1045 AAGTCAGGCCCTGGTCT 1104

## RESULT 5

AAS01706/c

ID AAS01706 standard; RNA; 1131 BP.

XX AAS01706;

XX 29-AUG-2001 (first entry)

XX TADG-16 antisense RNA transcript.

XX Human; extracellular serine protease; tumour antigen derived gene-16;  
KW TADG-16; ovarian carcinoma; breast cancer; lung cancer; colon cancer;  
KW prostate cancer; HLA type; ss.  
XX Homo sapiens.

XX W0200127257-A1.

XX 19-APR-2001.

XX 13-OCT-2000; 2000WO-US28558.

XX 14-OCT-1999; 99US-0418527.

XX (UYAR-) UNIV ARKANSAS.

XX O'Brien TJ, Underwood LJ, Shigemasa K;

XX WPI; 2001-273769/28.

XX New tumour antigen-derived gene-16 protein, useful for diagnosis and  
PT treatment of ovarian, breast, lung, colon and prostate cancer -  
XX Example 7; Page 85; 124pp; English.

XX The present sequence is for human TADG-16 antisense RNA transcript.

CC Tumour antigen derived gene-16 protein, TADG-16 (AAU02223), is a novel  
CC human extracellular serine protease. TADG-16 is expressed in normal  
CC ovaries and testes and in certain ovarian carcinomas. TADG-16 contains  
CC the conserved catalytic triad, His-Asp-Ser, and a signal secretion  
CC sequence characteristic of the serine protease family. An antisense  
CC oligonucleotide having a complementary sequence to the TADG-16 nucleic  
CC acid is useful for treating various cancers, including ovarian, breast,  
CC lung, colon and prostate. The TADG-16 nucleic acid, TADG-16 protein and  
CC antibodies specific to TADG-16 are useful for the diagnosis of cancer.  
CC TADG-16 protein or its fragments are useful for vaccinating an  
CC individual against TADG-16. Numerous TADG-16 peptides (AAU02225-AAU02384)  
CC are tested for their binding affinity to the 8 haplotypes HLA A0201,  
CC HLA A0205, HLA A1, HLA A24, HLA B7, HLA B8, HLA B2702, and HLA B4403.  
XX

SQ Sequence 1131 BP; 242 A; 340 C; 332 G; 217 U; 0 other;

Query Match 98.4%;

Best Local Similarity 99.6%; Pred. No. 2.1e-248;

Matches 1076; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 AGGAGGAGAGGGGCGCTCAGGCGCGGAGAGAGCCATGGCGCGCGGCGCGCTG 60  
|||||

Db 1107 AGGAGGACAGGGGGGCGTCAAGCCGCGGAGAGGAGCCATGGCGCGCGGGGCGCTG 1048  
Qy 61 CTGCTGGCGTGTCTGCTGGCTCGGGCTGGACTCAGGAGCGGAGTCCGAGGAGCGCGC 120  
Db 1047 CTGCTGGCGTGTCTGCTGGCTCGGGCTGGACTCAGGAGCGGAGTCCGAGGAGCGCGC 988  
Qy 121 CCCTTATCAGGACCATGCGGCCGACGGGTGATCAGCTCGCGCATCGTGGGTGAGAGGAC 180  
Db 987 CCGTTATCAGGACCATGCGGCCGACGGGTGATCAGCTCGCGCATCGTGGGTGAGAGGAC 928  
Qy 181 GCCGAATCGGGCGTGGCGGTGGAGGAGCGCTCGCGCTGTGGATTCCCACTATGC 240  
Db 927 GCCGAATCGGGCGTGGCGGTGGAGGAGCGCTCGCGCTGTGGATTCCCACTATGC 868  
Qy 241 GGAGTGGAGCTGTCTAGCCAGCGCTGGGCACTCACGGCGGGGCACTGCTTTGAAACCTAT 300  
Db 867 GGAGTGGAGCTGTCTAGCCAGCGCTGGGCACTCACGGCGGGGCACTGCTTTGAAACCTAT 808  
Qy 301 AGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA 360  
Db 807 AGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA 748  
Qy 361 TCCCTTCTGGAGCTCGAGGCTTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGC 420  
Db 747 TCCCTTCTGGAGCTCGAGGCTTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGC 688  
Qy 421 CUTCCTACTCTGGGAAATTCACCTTATGACATTTGCTTGGTGAAGCTCTGCACTCTGC 480  
Db 687 CUTCCTACTCTGGGAAATTCACCTTATGACATTTGCTTGGTGAAGCTCTGCACTCTGC 628  
Qy 481 ACCTACACTAAACACATCCAGCCATCTCTCCAGGCTCCACATTTGAGTTTGAAGAAC 540  
Db 627 ACCTACACTAAACACATCCAGCCATCTCTCCAGGCTCCACATTTGAGTTTGAAGAAC 568  
Qy 541 CGGACAGACTCTGGGTGACTGGCTGGGGTACATCAAGAGGATGAGGCACTGCCATCT 600  
Db 567 CGGACAGACTCTGGGTGACTGGCTGGGGTACATCAAGAGGATGAGGCACTGCCATCT 508  
Qy 601 CCCACACCTTCCAGGAAATTCAGTCCGATCATCAACACTATGCTGCTGCAACCACTC 660  
Db 507 CCCACACCTTCCAGGAAATTCAGTCCGATCATCAACACTATGCTGCTGCAACCACTC 448  
Qy 661 TTCTCAAGTACAGTTTCCGAAAGGACATCTTTGGAGACATGTTTGTGCTGCAATGCC 720  
Db 447 TTCTCAAGTACAGTTTCCGAAAGGACATCTTTGGAGACATGTTTGTGCTGCAATGCC 388  
Qy 721 CAAGGCGGGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Db 387 CAAGGCGGGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 328  
Qy 781 GGAGTGTGTATCAGATTGGAGTCTGAGTGGGAGTGGGCTGGTGGGCTGCTGCTGCTGCTGCT 840  
Db 327 GGAGTGTGTATCAGATTGGAGTCTGAGTGGGAGTGGGCTGGTGGGCTGCTGCTGCTGCTGCT 268  
Qy 841 CCGGTGTGTATCAGATTGGAGTCTGAGTGGGAGTGGGCTGGTGGGCTGCTGCTGCTGCTGCTGCT 900  
Db 267 CCGGTGTGTATCAGATTGGAGTCTGAGTGGGAGTGGGCTGGTGGGCTGCTGCTGCTGCTGCTGCT 208  
Qy 901 AGTGCACTGCTCCAGGCAACCTTCTGCGGCACTTCTTTTCCCTTCTTCTGCTGCTGCTGCTGCT 960  
Db 207 AGTGCACTGCTCCAGGCAACCTTCTGCGGCACTTCTTTTCCCTTCTTCTGCTGCTGCTGCTGCT 148  
Qy 961 CTCCCACTCTCTGGGCGGCTGCTGAGCTTACCTGAGCCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1019  
Db 147 CTCCCACTCTCTGGGCGGCTGCTGAGCTTACCTGAGCCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 88  
Qy 1020 AGTGCACTGCT 1079  
Db 87 AAGTCAAGCCCTGCT 28

RESULT 6  
AA52259

AA52259 standard; DNA; 1100 BP.

AA52259;

25-JUN-1999 (first entry)

Protein PRO303 cDNA clone DNA42551-1217.

Secreted protein; transmembrane protein; human; enterocolitis;  
Zollinger-Ellison syndrome; gastrointestinal ulceration;  
congenital microvillus atrophy; skin disease; cell growth;  
abnormal keratinocyte differentiation; psoriasis; epithelial cancer;  
Parkinson's disease; Alzheimer's disease; ALS; neuropathy;  
fibromodulin; dermal scarring; Usher Syndrome; Atrophica areata;  
anti-thrombotic; wound healing; tissue repair; ss.

Homo sapiens.

WO914328-A2.

25-MAR-1999.

16-SEP-1998; 98WO-US19330.

25-NOV-1997; 97US-0066840.

17-SEP-1997; 97US-0059113.

17-SEP-1997; 97US-0059115.

17-SEP-1997; 97US-0059117.

17-SEP-1997; 97US-0059119.

17-SEP-1997; 97US-0059121.

17-SEP-1997; 97US-0059122.

17-SEP-1997; 97US-0059184.

18-SEP-1997; 97US-0059263.

18-SEP-1997; 97US-0059266.

15-OCT-1997; 97US-0062125.

17-OCT-1997; 97US-0062285.

17-OCT-1997; 97US-0062287.

21-OCT-1997; 97US-0063486.

24-OCT-1997; 97US-0062814.

24-OCT-1997; 97US-0062816.

24-OCT-1997; 97US-0063045.

24-OCT-1997; 97US-0063120.

24-OCT-1997; 97US-0063121.

24-OCT-1997; 97US-0063127.

27-OCT-1997; 97US-0063128.

27-OCT-1997; 97US-0063329.

28-OCT-1997; 97US-0063541.

28-OCT-1997; 97US-0063542.

28-OCT-1997; 97US-0063544.

28-OCT-1997; 97US-0063549.

28-OCT-1997; 97US-0063550.

28-OCT-1997; 97US-0063554.

29-OCT-1997; 97US-0063435.

29-OCT-1997; 97US-0063704.

29-OCT-1997; 97US-0063732.

29-OCT-1997; 97US-0063738.

29-OCT-1997; 97US-0063734.

29-OCT-1997; 97US-0064215.

29-OCT-1997; 97US-0063735.

31-OCT-1997; 97US-0063870.

31-OCT-1997; 97US-0064103.

03-NOV-1997; 97US-0064248.

07-NOV-1997; 97US-0065186.

12-NOV-1997; 97US-0065186.

17-NOV-1997; 97US-0065846.

18-NOV-1997; 97US-0065846.

21-NOV-1997; 97US-0066120.

21-NOV-1997; 97US-0066120.

21-NOV-1997; 97US-0066364.

24-NOV-1997; 97US-0066772.

24-NOV-1997; 97US-0066466.

24-NOV-1997; 97US-0066770.

24-NOV-1997; 97US-0066511.

24-NOV-1997; 97US-0066453.  
(GETH ) GENENTECH INC.  
Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;  
WPI; 1999-229533/19.  
P-PSDB; AAY13388.  
New isolated human genes and polypeptides used in, e.g. treatment of  
gastrointestinal ulceration  
Claim 2; Fig 91; 320pp; English.  
AA52213-74 encode secreted and transmembrane human proteins, and are  
obtained from cDNA libraries, prepared from fetal lung, fetal kidney,  
fetal brain, fetal liver and fetal retina. The encoded polypeptides  
have specific uses based on their homology to known polypeptides,  
e.g. PRO211 and PRO217 can be used for disorders associated with the  
preservation and maintenance of gastrointestinal mucosa and the repair  
of acute and chronic mucosal lesions (e.g. enterocolitis,  
Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital  
microvillus atrophy), skin diseases associated with abnormal  
keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as  
lung squamous cell carcinoma of the vulva and gliomas), potent effects on  
cell growth and development, diseases related to growth or survival of  
nerve cells including Parkinson's disease, Alzheimer's disease, ALS,  
neuropathies or cancer. PRO265 can be used as for fibromodulin, e.g. for  
reducing dermal scarring. PRO264 can be used as a target for anti-tumor  
drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophia  
creata; PRO269 can be used as an anti-thrombotic agent; PRO287  
polypeptides and portions may have therapeutic applications in wound  
healing and tissue repair; PRO317 can be used for treating problems of  
the kidney, uterus, endometrium, blood vessels, or related tissue, e.g.  
in the heart of genital tract.  
SQ Sequence 1100 BP; 225 A; 321 C; 314 G; 240 T; 0 other;  
Query Match 97.7%; Score 1055.8; DB 20; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 1.3e-246;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 21 GCGCGGGAGAGGAGCCATGGCGCGCGGGCGCGTGTCTGTGGCGCTGCTGCTGGC 80  
Db 1 GCGCGGGAGAGGAGGAGCCATGGCGCGCGGGCGCGTGTCTGTGGCGCTGCTGCTGGC 60  
Qy 81 TCGGCTGACTCAGGAAGCGGAGTCCAGAGGCGCGCGCTTATCAGGACCATGGG 140  
Db 61 TCGGCTGGACTCAGGAAGCGGAGTCCAGAGGCGCGCGCTTATCAGGACCATGGG 120  
Qy 141 CCGACGGGTATCATCGTCGCGCATCGTGGTGGAGAGGACGCGCAACTCGGGCGTGGCC 200  
Db 121 CCGACGGGTATCATCGTCGCGCATCGTGGTGGAGAGGACGCGCAACTCGGGCGTGGCC 180  
Qy 201 GTGCAGGGAGCGCTGGCGCTGTGGATGCCAGTATCCAGATGAGCCCTGCTCAGCCA 260  
Db 181 GTGCAGGGAGCGCTGGCGCTGTGGATGCCAGTATCCAGATGAGCCCTGCTCAGCCA 240  
Qy 261 CCGTGGGCATCAGCGGGCGGCACTGCTTTGAACCTATAGTAGACCTTAGTATCCCTC 320  
Db 241 CCGTGGGCATCAGCGGGCGGCACTGCTTTGAACCTATAGTAGACCTTAGTATCCCTC 300  
Qy 321 CCGGTGGATGTTCCAGTTTGGCCAGCTGACTTCCATGCGCATCTTCTGTGAGCCTCAGGC 380  
Db 301 CCGGTGGATGTTCCAGTTTGGCCAGCTGACTTCCATGCGCATCTTCTGTGAGCCTCAGGC 360  
Qy 381 CTACTACACCGGTACTTCTGATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 440  
Db 361 CTACTACACCGGTACTTCTGATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 420  
Qy 441 ACCCTATGACATTCGCTGGTGAAGCTGTCTGACCTGTACCTACATAACACATCCA 500  
Db 421 ACCCTATGACATTCGCTGGTGAAGCTGTCTGACCTGTACCTACATAACACATCCA 480

Qy 501 GCCCATCTGTCTCCAGGCTCCACATTTTGTAGTTTGTAGAACCGGACAGACTGCTGGTGAC 560  
Db 481 GCCCATCTGTCTCCAGGCTCCACATTTTGTAGTTTGTAGAACCGGACAGACTGCTGGTGAC 540  
Qy 561 TGGCTGGGGGTACATCAAGAGGATGAGGACTGCGCATCTCCCAACACCTCCAGGAAGT 620  
Db 541 TGGCTGGGGGTACATCAAGAGGATGAGGACTGCGCATCTCCCAACACCTCCAGGAAGT 600  
Qy 621 TCAGTGGCCATCATAAACAACCTCTATGTGCAACACCTCTTCTCCTCAAGTACAGTTCCG 680  
Db 601 TCAGTGGCCATCATAAACAACCTCTATGTGCAACACCTCTTCTCCTCAAGTACAGTTCCG 660  
Qy 681 CAAGGACATCTTTGGAGCATGTTTGTGCTGCAATGCCAAGCGGGAAGATGCCCTG 740  
Db 661 CAAGGACATCTTTGGAGCATGTTTGTGCTGCAACGCCAAGCGGGAAGATGCCCTG 720  
Qy 741 CTTCCGCTGACTCAGTGGACCTTGGCCTGTAAAGAAATGAGTGTGTATCAGATGG 800  
Db 721 CTTCCGCTGACTCAGTGGACCTTGGCCTGTAAAGAAATGAGTGTGTATCAGATGG 780  
Qy 801 AGTCGTGAGCTGGGAGTGGGCTGTGTCGGCCCAATCGGCCGTGTCTACACCAATAT 860  
Db 781 AGTCGTGAGCTGGGAGTGGGCTGTGTCGGCCCAATCGGCCGTGTCTACACCAATAT 840  
Qy 861 CAGCCACCATCTTGTAGTGATCCAGAAGCTGATGCGCCAGAGTGGCATGTCCAGCCAGA 920  
Db 841 CAGCCACCATCTTGTAGTGATCCAGAAGCTGATGCGCCAGAGTGGCATGTCCAGCCAGA 900  
Qy 921 CCCCTCTCGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACATCTCGGGCGCGT 980  
Db 901 CCCCTCTCGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACATCTCGGGCGCGT 960  
Qy 981 CTGAGCCTACCTGAGCCCATGCGCTGGGCGCACTGCCAAGTCAAGCCTGAGCCCTGTTCTCT 1040  
Db 961 CTGAGCCTACCTGAGCCCATGCGCTGGGCGCACTGCCAAGTCAAGCCTGAGCCCTGTTCTCT 1020  
Qy 1041 CTGCTTGTGTGTAATAACACATTCAGTTGATGCCT 1079  
Db 1021 CTGCTTGTGTGTAATAACACATTCAGTTGATGCCT 1059  
RESULT 7  
AA72417  
ID AAF72417 standard; cDNA; 1100 BP.  
XX  
AC AAF72417;  
XX  
DT 24-APR-2001 (first entry)  
XX  
DE Human PRO303 cDNA.  
XX  
KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;  
antiparkinsonian nootropic; neuroprotective; vulnery; cardiac;  
antiangiogenic; vasotropic; antihistaminic; antirheumatic; cancer;  
antiarthritic; antinfertility; antidiabetic; antiviral; diabetes;  
ophthalmological; gene therapy; skin disease; gastrointestinal disorder;  
ischaemia; inflammation; ss.  
OS Homo sapient.  
PN WO200104311-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 22-FEB-2000; 2000WO-US04414.  
XX  
PR 07-JUL-1999; 99US-0143048.  
PR 26-JUL-1999; 99US-0145698.  
PR 28-JUL-1999; 99US-0146222.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 05-JAN-2000; 99WO-US00219.  
XX (GETH ) GENENTECH INC.  
PA  
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;  
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kijavini IJ;  
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;  
PI Williams PM, Wood WI;  
XX WPI: 2001-081051/09.  
DR P-PSDB; AAB80256.  
XX  
XX Sixty one nucleic acids encoding PRO polypeptides which are useful in  
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung  
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.  
PT Alzheimer's disease) -  
XX  
XX Claim 2; Fig 91; 393pp; English.  
XX  
XX The present sequence is one of sixty one nucleic acids encoding novel  
CC secreted and transmembrane PRO polypeptides. The PRO polypeptides are  
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung  
CC squamous cell carcinoma), gastrointestinal disorders (e.g.  
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.  
CC ischaemia, atherosclerosis), inflammatory disorders such as coronary  
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and  
CC diabetes and retinal disorders such as retinitis pigmentosum.  
CC The PRO nucleic acids have applications in molecular biology, including  
CC use as hybridization probes, and in chromosome and gene mapping.  
XX  
XX Sequence 1100 BP; 225 A; 321 C; 314 G; 240 T; 0 other;  
Query Match 97.7%; Score 1055.8; DB 22; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 1.3e-246;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 21 GGCGCGGGAGAGAGGCCATGGCGCGCGCGCGCGCTGCTGCTGCGCGTGTGCTGGC 80  
DB 1 GGCGCGGGAGAGAGGCCATGGCGCGCGCGCGCGCTGCTGCTGCGCGTGTGCTGGC 60  
QY 81 TCGGGTGGACTCAGGAAGCCGAGTGCAGAGAGCGCGCGCTTATCAGGACCATGCGG 140  
DB 61 TCGGGTGGACTCAGGAAGCCGAGTGCAGAGAGCGCGCGCTTATCAGGACCATGCGG 120  
QY 141 CCACGGGTATCATCAGCTCGCGCATCGTGGTGAGAGAGCGCGAACTCGGCGTTGGCC 200  
DB 121 CCACGGGTATCATCAGCTCGCGCATCGTGGTGAGAGAGCGCGAACTCGGCGTTGGCC 180  
QY 201 GTGGCAGGGAGCGCTCGCGCTGTGGGATTCGCCAGTATCGGAGTGCAGCTGTAGCCA 260  
DB 181 GTGGCAGGGAGCGCTCGCGCTGTGGGATTCGCCAGTATCGGAGTGCAGCTGTAGCCA 240  
QY 261 CGCTGGGCATCATCGGCGCGCATCGCTTTGAAACCTATAGTACCTTAGTATCCCTC 320  
DB 241 CGCTGGGCATCATCGGCGCGCATCGCTTTGAAACCTATAGTACCTTAGTATCCCTC 300  
QY 321 CGGTTGGATGGTCCAGTTGGCGAGTGCAGTCCATGCCATCCTTCGTAGCGCTCAGGC 380  
DB 301 CGGTTGGATGGTCCAGTTGGCGAGTGCAGTCCATGCCATCCTTCGTAGCGCTCAGGC 360  
QY 381 CTACTACACCCGTTACTTCTGATCAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 440  
DB 361 CTACTACACCCGTTACTTCTGATCAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 420

QY 441 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGCACTGTACCTACCTACCTAAACATCCA 500  
DB 421 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGCACTGTACCTACCTAAACATCCA 480  
QY 501 GCCCATCTGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGAGAGTGTGGGTGAC 560  
DB 481 GCCCATCTGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGAGAGTGTGGGTGAC 540  
QY 561 TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCTCCACCTCCACACCTCCAGGAAT 620  
DB 541 TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCTCCACCTCCACACCTCCAGGAAT 600  
QY 621 TCAGTCCGCTCATATAAACAACCTATGTGCAACCACTCTCTCTCAAGTACAGTTTCG 680  
DB 601 TCAGTCCGCTCATATAAACAACCTATGTGCAACCACTCTCTCTCAAGTACAGTTTCG 660  
QY 681 CAAGGACATCTTGGAGACATGGTTGTGCTGGCAATGCCAAGCGGGAAGGATGCCTG 740  
DB 661 CAAGGACATCTTGGAGACATGGTTGTGCTGGCAATGCCAAGCGGGAAGGATGCCTG 720  
QY 741 CTTGGTCACTCAGTGGAGCCCTTGGCTGTAAACAAGATGAGTGTGATCAGATTGG 800  
DB 721 CTTGGTCACTCAGTGGAGCCCTTGGCTGTAAACAAGATGAGTGTGATCAGATTGG 780  
QY 801 AGTGTGAGTGGGAGTGGCTGTGCTGGCCCAATGCGGCGCTGTCTACACCAATAT 860  
DB 781 AGTGTGAGTGGGAGTGGCTGTGCTGGCCCAATGCGGCGCTGTCTACACCAATAT 840  
QY 861 CAGCCACCATCTTGGAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGCCAGCCAGA 920  
DB 841 CAGCCACCATCTTGGAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGCCAGCCAGA 900  
QY 921 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGCGCGT 980  
DB 901 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGCGCGT 960  
QY 981 CTGAGCTACTCTGAGCCCATGAGCCCTGGGCGCACTGCCAAGTCCAGCCCTGTCTCTT 1040  
DB 961 CTGAGCTACTCTGAGCCCATGAGCCCTGGGCGCACTGCCAAGTCCAGCCCTGTCTCTT 1020  
QY 1041 CTGCTTGTGTTGTTAATAACACATTCAGTTGATGCT 1079  
DB 1021 CTGCTTGTGTTGTTAATAACACATTCAGTTGATGCT 1059  
XX AAD27418 standard; DNA; 1077 BP.  
XX AAD27418;  
XX 18-APR-2002 (first entry)  
XX Human eosinophil serine protease-1 (esp-1) like enzyme. DNA #3.  
XX Human; eosinophil serine protease-1; esp-1; enzyme; antiinflammatory;  
XX antiasthmatic; antiallergic; osteopathic; cytostatic; dermatological;  
XX asthma; airway allergy; chronic obstructive pulmonary disease; COPD;  
XX osteoporosis; dermatitis; Paget's disease; therapy; ds.  
XX Homo sapiens.  
XX WO200198503-A2.  
XX 27-DEC-2001.  
XX 20-JUN-2001; 2001WO-EP06936.  
XX 21-JUN-2000; 2000US-212844P.  
XX 31-OCT-2000; 2000US-24417P.  
XX 30-MAR-2001; 2001US-279766P.  
XX

(FARB ) BAYER AG.  
Xiao Y;  
WPI; 2002-122283/16.  
Novel purified human eosinophil serine protease 1-like enzyme, useful for identifying modulators of enzyme activity for treating Paget's disease, osteoporosis, allergy, asthma  
Disclosure; Fig 7; 131pp; English.  
The invention relates to a purified human eosinophil serine protease-1 (esp-1) like enzyme. Esp-1 like enzyme is useful in in-vitro or in-vivo assays to identify test compounds with potential therapeutic or diagnostic value. Esp-1 like enzyme modulator is useful for treating esp-1 like enzyme dysfunction related diseases condition such as asthma, allergy, chronic obstructive pulmonary disease (COPD) or osteoporosis. Esp-1 like enzyme is also useful in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases related to presence of mutations in the nucleic acid sequences which encode the enzyme. Pharmaceutical composition comprising esp-1 like enzyme is useful for treating dermatitis, Paget's disease, and preventing degradation of bone implants particularly dental implants. The present sequence is human esp-1 like enzyme DNA.  
Sequence 1077 BP; 202 A; 320 C; 314 G; 241 T; 0 other;  
Query Match 97.6%; Score 1054.8; DB 24; Length 1077;  
Best Local Similarity 99.8%; Pred. No. 2.3e-246;  
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 22 GCCGCGGAGAGGAGGATGGCGCGCGGGCGTGTCTGGCGCTGCTGCTGGCT 81  
DB 1 GCCGCGGAGAGGAGGAGGATGGCGCGCGGGCGTGTCTGGCGCTGCTGCTGGCT 60  
QY 82 CGGGCTGGACTCAGGAAGCGGAGTGCAGAGGCGCGCGCTTATCAGGACCATCGGC 141  
DB 61 CGGGCTGGACTCAGGAAGCGGAGTGCAGAGGCGCGCGCTTATCAGGACCATCGGC 120  
QY 142 CGAGGGTATCAGTCCGATCGTGGTGGAGAGCGCGGAGTGCAGGACCATCGGC 201  
DB 121 CGAGGGTATCAGTCCGATCGTGGTGGAGAGCGCGGAGTGCAGGACCATCGGC 180  
QY 202 TGGCAGGGAGCGCTGCGCTGTGGATTCACATGCGGAGTGCAGGACCATCGGC 261  
DB 181 TGGCAGGGAGCGCTGCGCTGTGGATTCACATGCGGAGTGCAGGACCATCGGC 240  
QY 262 CGCTGGGACATCAGCGCGCGGCTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCC 321  
DB 241 CGCTGGGACATCAGCGCGCGGCTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCC 300  
QY 322 GGGTGGATGTCAGTTTGGCCAGCTGACATCCATGCCATCCTTCTGGAGCTGCAGGCC 381  
DB 301 GGGTGGATGTCAGTTTGGCCAGCTGACATCCATGCCATCCTTCTGGAGCTGCAGGCC 360  
QY 382 TACTACACCGTACTTCGATCGAATATCTATCTAGCGCTCGCTACCTGGGGAATCA 441  
DB 361 TACTACACCGTACTTCGATCGAATATCTATCTAGCGCTCGCTACCTGGGGAATCA 420  
QY 442 CCCTATGACATTCGCTTGGTGAAGCTGCTGCACCTGTACCTACATTAACACATCCAG 501  
DB 421 CCCTATGACATTCGCTTGGTGAAGCTGCTGCACCTGTACCTACATTAACACATCCAG 480  
QY 502 CCATATGCTTCAGGCGCTCCACATTTAGTTTGGAGAACCGGACAGACTGCTGGGTGACT 561  
DB 481 CCATATGCTTCAGGCGCTCCACATTTAGTTTGGAGAACCGGACAGACTGCTGGGTGACT 540  
QY 562 GGGTGGGGTACATCAAGAGGATGAGGACCTGCCATCTCCCGACACCTCCAGGAGTT 621  
DB 541 GGGTGGGGTACATCAAGAGGATGAGGACCTGCCATCTCCCGACACCTCCAGGAGTT 600  
QY 622 CAGGTCGCCATCAATAAACAACCTATATGTGCAACCAACCTCTTCCTCAAGTACAGTTTCCGC 681

DB 601 CAGGTCCCATCATATAACAACCTCTATGTGCAACCAACCTTCTCTCAAGTACAGTTTCCGC 660  
QY 682 AAGGACATCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGCGGGAGGATGCTGC 741  
DB 661 AAGGACATCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGCGGGAGGATGCTGC 720  
QY 742 TTCGGTCACTCAGGTGAGCCTTGGCTGTAAACAAGATGAGCTGTGTATCAGATTGGA 801  
DB 721 TTCGGTCACTCAGGTGAGCCTTGGCTGTAAACAAGATGAGCTGTGTATCAGATTGGA 780  
QY 802 GTCTGTAGCTGGGAGTGGGCTGTGCTGGCCCAATCGGCCCGTGTCTACACAATATC 861  
DB 781 GTCTGTAGCTGGGAGTGGGCTGTGCTGGCCCAATCGGCCCGTGTCTACACAATATC 840  
QY 862 AGCCACCACTTTGAGTGGATCCAGAACTGATGCCCCAGAGTGGCATGTCCACCCAGAC 921  
DB 841 AGCCACCACTTTGAGTGGATCCAGAACTGATGCCCCAGAGTGGCATGTCCACCCAGAC 900  
QY 922 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGCGCTCTCCACTCTCTGGGCGCGTC 981  
DB 901 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGCGCTCTCCACTCTCTGGGCGCGTC 960  
QY 982 TGAGCCTTACCTGAGCCCATGCGCTGGGCGCACTGCCAAGTCAAGGCCCTGTCTCTTC 1041  
DB 961 TGAGCCTTACCTGAGCCCATGCGCTGGGCGCACTGCCAAGTCAAGGCCCTGTCTCTTC 1020  
QY 1042 TGTCTTGTGGTAAATAAACACATTCAGTTGATGCCT 1079  
DB 1021 TGTCTTGTGGTAAATAAACACATTCAGTTGATGCCT 1058  
RESULT 9  
AAZ45673/C  
ID -AAZ45673 standard; DNA; 1079 BP.  
XX  
AC AAZ45673;  
XX AC  
XX AC  
DT 06-APR-2000 (first entry)  
XX  
DE Nucleotide sequence of prostate specific gene clone 4050626.  
XX  
KW Human; prostate specific gene; PSG; prostate tumour; prostate cancer; metastasis; prostate; ss.  
XX Homo sapiens.  
XX  
PN WO9960162-A1.  
XX  
PD 25-NOV-1999.  
XX  
PF 12-MAY-1999; 99WO-US10548.  
XX  
PR 21-MAY-1998; 98US-0086265.  
XX  
PA (DIAD-) DIADEXUS LLC.  
XX  
PI Ali S, Salceda S, Sun Y, Cafferkey R;  
XX  
DR WPI; 2000-126384/11.  
XX  
PT Diagnosing, monitoring, and staging prostate cancer  
XX  
PS Example 1; Page 29; 30pp; English.  
XX  
CC The present sequence represents a human prostate specific gene (PSG) (gene ID 236019). The mRNA levels for PSG are about 20 fold higher in the prostate than in other tissues. The PSG mRNA was found to be overexpressed in prostate tumour samples. The specification describes a method for diagnosing the presence of, or metastatic potential of, prostate cancer in a patient. The method comprises measuring PSG levels in a cell, tissue or bodily fluid sample of the patient and a control (i.e. a normal human without cancer), where increased PSG levels in



CC the patient compared to the control is associated with the presence of,  
 CC or metastasis of prostate cancer. The method can also be used for  
 CC staging prostate cancer in a patient. The methods are used to detect,  
 CC monitor, stage and give a prognosis for prostate cancer.  
 XX  
 SQ

Query Match 97.4%; Score 1052.4; DB 21; Length 1079;  
 Best Local Similarity 99.6%; Pred. No. 8.7e-246;  
 Matches 1053; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 23 CCGCGGAGAGAGAGCCATGGCGCGCGCGCGCGCGCTGCTGGCGCTGCTGGCGCTC 82  
 DB 1077 CCGCGGAGAGAGAGCCATGGCGCGCGCGCGCGCTGCTGGCGCTGCTGGCGCTC 1018  
 QY 83 GGGCTGGACTCAGGAAGCGGAGTGCAGAGAGCGCGCTTATCAGGACCATGCGGCC 142  
 DB 1017 GGGCTGGACTCAGGAAGCGGAGTGCAGAGAGCGCGCTTATCAGGACCATGCGGCC 958  
 QY 143 GACGGGTATCATCTGCGGCTGCTGGGTGAGAGGACGCGGAACCTGCGGCGTTGGCGCT 202  
 DB 957 GACGGGTATCATCTGCGGCTGCTGGGTGAGAGGACGCGGAACCTGCGGCGTTGGCGCT 898  
 QY 203 GGCAGGGAGCGCTGCGCTGTTGGATTCGCCACCTATCGGAGTGCAGTGCAGCCACC 262  
 DB 897 GGCAGGGAGCGCTGCGCTGTTGGATTCGCCACCTATCGGAGTGCAGTGCAGCCACC 838  
 QY 263 GCTGGGCACTACGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCG 322  
 DB 837 GCTGGGCACTACGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCG 778  
 QY 323 GGTGGATGCTCAGTTGGCCAGCTGACTTCCATGCGCATCTTCTGGAGCGCTCAGCGCT 382  
 DB 777 GGTGGATGCTCAGTTGGCCAGCTGACTTCCATGCGCATCTTCTGGAGCGCTCAGCGCT 718  
 QY 383 ACTACACCGCTTACTTCGTATCAATATCTATGAGCGCTCGCTACCTGGGGAATTCAC 442  
 DB 717 ACTACACCGCTTACTTCGTATCAATATCTATGAGCGCTCGCTACCTGGGGAATTCAC 658  
 QY 443 CCTATGACATGCTGCTGGTGAAGCTGCTGCACTGCTGCACTACACTAAACACATCCAGC 502  
 DB 657 CCTATGACATGCTGCTGGTGAAGCTGCTGCACTGCTGCACTACACTAAACACATCCAGC 598  
 QY 503 CCATCTGCTCCAGGCTCCCATTTGAGTTGAGAACCGGAGACTGCTGGTGAAGT 562  
 DB 597 CCATCTGCTCCAGGCTCCCATTTGAGTTGAGAACCGGAGACTGCTGGTGAAGT 538  
 QY 563 GCTGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCACACCTCCAGGAGTTC 622  
 DB 537 GCTGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCACACCTCCAGGAGTTC 478  
 QY 623 AGGTGCCATCAATCAACACTATGTGCAACCACTCTTCTCAAGTACAGTTCCGCA 682  
 DB 477 AGGTGCCATCAATCAACACTATGTGCAACCACTCTTCTCAAGTACAGTTCCGCA 418  
 QY 683 AGGACATCTTTGGACATGTTTGTGTCGCAATGCCAGCGGGAAGGATGCTGCT 742  
 DB 417 AGGACATCTTTGGACATGTTTGTGTCGCAATGCCAGCGGGAAGGATGCTGCT 358  
 QY 743 TCGGTGACTCAGGTGAGCCCTTGGCTGTAAACAAGATGAGTGTGATCAGATTGAG 802  
 DB 357 TCGGTGACTCAGGTGAGCCCTTGGCTGTAAACAAGATGAGTGTGATCAGATTGAG 298  
 QY 803 TCGTGAAGTGGGAGTGGCTGTGGTGGCGCCCAATCGCGCCGGTCTTACACCAATATCA 862  
 DB 297 TCGTGAAGTGGGAGTGGCTGTGGTGGCGCCCAATCGCGCCGGTCTTACACCAATATCA 238  
 QY 863 GCCACCACTTTGAGTGCATCAGAGCTGATGGCCAGAGTGGATGTCCAGCAGACC 922  
 DB 237 GCCACCACTTTGAGTGCATCAGAGCTGATGGCCAGAGTGGATGTCCAGCAGACC 178  
 QY 923 CCTCCTGGCCACTACTCTTTTCTCTCTCTGGGCTCTCCCACTCCTGGGCGCGTCT 982  
 DB 923 CCTCCTGGCCACTACTCTTTTCTCTCTCTGGGCTCTCCCACTCCTGGGCGCGTCT

DB 177 CCTCTGCGCCACTACTCTTTTCTCTCTCTGGGCTCTCCCACTCTCTGGGCGCGTCT 118  
 QY 983 GAGCCTACCTGAGCCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1042  
 DB 117 GAGCCTACCTGAGCCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 58  
 QY 1043 GTCCTGTTGGTGTATAAACAACATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1079  
 DB 57 GTCCTGTTGGTGTATAAACAACATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 21

## RESULT 10

AAV59119  
 ID AAV59119 standard; DNA; 1100 BP.

XX AAV59119;

XX 07-JAN-1999 (first entry)

XX Nucleotide sequence of long isoform of HELA2.

XX Serine protease; regulation; cell activity; viability; HELA2; ATC2;  
 KW BCOM3; testis; fertility; suppressor; testicular germ cell cancer;  
 KW seminoma; testis-specific expression; antitumour; sperm development;  
 XX infertility; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FD 17..961

FT CDS

FT \*tag= a

FT /product= HELA2

PN W09836054-A1.

XX 20-AUG-1998.

XX 13-FEB-1998; 98WO-AU000085.

XX 18-NOV-1997; 97AU-0000422.

PR 13-FEB-1997; 97AU-0005101.

XX (AMRA-) AMRAD OPERATIONS PTY LTD.

XX Antalis TM, Hooper JD;

XX WPI; 1998-480768/41.

DR P-PSDB; AAW77297.

XX New serine protease(s) and kinase involved in regulating cell  
 activity and viability - particularly the testis-specific protease  
 HELA2 used for modulation of fertility and as tumour suppressor

Claim 6; Pages 62-64; 167pp; English.

The present sequence represents the nucleotide sequence of the long  
 isoform of HELA2. cDNA generated from HELA cells and PAI-2 expressing  
 HELA cells was amplified using PCR primers AAV48312-13. Three new  
 sequences were detected in the 480 bp amplicon. These sequences are  
 designated HELA2 and ATC2 which have high homology to serine proteases  
 and BCOM3 which has homology to a kinase. The proteins are involved in  
 or associated with regulation of cell activity and/or viability.  
 Administration of recombinant HELA2 (also called testisin) is used to  
 increase fertility. Downregulation of HELA2 reduces fertility. HELA2 is  
 also a suppressor of testicular germ cell cancers (seminoma) and is also  
 expressed in some non-testicular cancers (of colon, pancreas, prostate  
 and ovary), so is a marker/potential therapeutic target for cancer. The  
 promoter from the HELA2 gene is useful for testis-specific expression of  
 other genes, e.g. for gene therapy or modulation of fertility. Drugs  
 that block activity of HELA2 should have antitumour activity (other than  
 in testis) while in testis recombinant HELA2 should stop growth of  
 tumours and normalise sperm development (eliminating the need for  
 orchidectomy). Identification of mutant forms of HELA2 can be used to

CC	diagnose infertility.
XX	
SQ	Sequence 1100 BP; 226 A; 319 C; 313 G; 242 T; 0 other;
	Query Match 97.18; Score 1049.6; DB 19; Length 1100;
	Best Local Similarity 99.68; Pred. No. 4.2e-245;
	Matches 1052; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	24 CGCGGGAGAGGCCATGCGCGCGCGCGCGCGCTGTCTGTGGCGCTGCTGCTGCCTCG 83
DB	1 CGCGGGAGAGGCCATGCGCGCGCGCGCGCGCTGTCTGTGGCGCTGCTGCTGCCTCG 60
QY	84 GGCTGGACTCAGGAAGCCGGAGTCGACAGAGCGCGCGCTTATCAGGACCATTGCGCGG 143
DB	61 GGCTGGACTCAGGAAGCCGGAGTCGACAGAGCGCGCGCTTATCAGGACCATTGCGCGG 120
QY	144 ACGGGTCTACATCGTCGCGCATCGTGCGGTGGAGAGGACCGCAACTCGGCGTTGGCCGG 203
DB	121 ACGGGTCTACATCGTCGCGCATCGTGCGGTGGAGAGGACCGCAACTCGGCGTTGGCCGG 180
QY	204 GCAGGGGAGCGCTGCGCCTGTGGGATTCACAGTATGCGGAGTAGCCCTGCTCAGCACCG 263
DB	181 GCAGGGGAGCGCTGCGCCTGTGGGATTCACAGTATGCGGAGTAGCCCTGCTCAGCACCG 240
QY	264 CTGGGCACTCAGCGCGGCGACTGTTTGAACCTATAGTGACCTTAGTATCCCTCCGG 323
DB	241 CTGGGCACTCAGCGCGGCGACTGTTTGAACCTATAGTGACCTTAGTATCCCTCCGG 300
QY	324 GTGGATGGTCAGTTTGGCCAGCTGACTTCCATGCCATCCCTTCTGGAGCCTCAGGCGCTA 383
DB	301 GTGGATGGTCAGTTTGGCCAGCTGACTTCCATGCCATCCCTTCTGGAGCCTCAGGCGCTA 360
QY	384 CTACACCGTTTACTTCTGATGAAATATCTATCTGAGCCCTGCTACCTGGGGAATTCACC 443
DB	361 CTACACCGTTTACTTCTGATGAAATATCTATCTGAGCCCTGCTACCTGGGGAATTCACC 420
QY	444 CTATGACATGCGCTTGGTGAAGCTGTGTCACCTGTACCTTACATAAACACATCCAGCC 503
DB	421 CTATGACATGCGCTTGGTGAAGCTGTGTCACCTGTACCTTACATAAACACATCCAGCC 480
QY	504 CATCTGTCTCAGGCTCCCATTTGAGTTTGGAGAACGGACAGACTGCTGGGTGACTGG 563
DB	481 CATCTGTCTCAGGCTCCCATTTGAGTTTGGAGAACGGACAGACTGCTGGGTGACTGG 540
QY	564 CTGGGGTACATCAAAGAGGATGAGGCACTGCCATCTCCCACACCCCTCCAGGAAGTTCA 623
DB	541 CTGGGGTACATCAAAGAGGATGAGGCACTGCCATCTCCCACACCCCTCCAGGAAGTTCA 600
QY	624 GGTGCGCATATAAACACTATGTGCAACACCTCTTCTCTCAAGTACAGTTTCCGCA 683
DB	601 GGTGCGCATATAAACACTATGTGCAACACCTCTTCTCTCAAGTACAGTTTCCGCA 660
QY	684 GGACATCTTGGACATGTTTGTGTCGCAAGTCCCAAGCGGAGGATGCTGCTT 743
DB	661 GGACATCTTGGACATGTTTGTGTCGCAAGTCCCAAGCGGAGGATGCTGCTT 720
QY	744 CGGTGACTCAGGTGAGCCCTTGGCTGTAAACAAGATGAGTGTGGTATCAGATTGGAGT 803
DB	721 CGGTGACTCAGGTGAGCCCTTGGCTGTAAACAAGATGAGTGTGGTATCAGATTGGAGT 780
QY	804 CGTGAGCTGGGAGTGGCTGTGTGCGCCCAATGCGGCGCGGTGCTACACCAATATCAG 863
DB	781 CGTGAGCTGGGAGTGGCTGTGTGCGGCGCAATGCGGCGCGGTGCTACACCAATATCAG 840
QY	864 CCACCACTTTCAGTGGATCCAGAGCTGATGCGCCAGAGTGGCATGTCCTCCAGCCAGACC 923
DB	841 CCACCACTTTCAGTGGATCCAGAGCTGATGCGCCAGAGTGGCATGTCCTCCAGCCAGACC 900
QY	924 CTCCTGCGCACTACTCTTTTCCCTCTCTCTGGCTCTCCACCTCCCTGGGCGCGGTCTG 983
DB	901 CTCCTGCGCGCTACTCTTTTCCCTCTCTCTGGCTCTCCACCTCCCTGGGCGCGGTCTG 960
QY	984 AGCCTACCTGAGGCCATCGACCTCGGGGCGCACTGCGAAGTCAGGCGCTGTTCTCTCTG 1043

Db	961	AGCCTACCTGAGGCCATCGACCTGGGGCCACCTGCCAAGTCAGGCCCTGCTCTCTCTG	1079
QY	1044	TCTTGTGTTGTAATAAACACATTCACATTCAGTTGATGCCT	1079
Db	1021	TCTTGTGTTGTAATAAACACATTCACATTCAGTTGATGCCT	1056
RESULT 11			
AAV59118	ID	AAV59118 standard; DNA; 1117 BP.	
XX	AAV59118;		
XX	DT	07-JAN-1999 (first entry)	
XX	Nucleotide sequence of short isoform of HELA2.		
XX	Serine protease; regulation; cell activity; viability; HELA2; ATC2;		
KW	BCOM3; testisin; fertility; suppressor; testicular germ cell cancer;		
KW	seminoma; testis-specific expression; antitumour; sperm development;		
KW	infertility; ss.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	CDS	19..963	
FT	/tag= a		
FT	/product= HELA2		
XX	WO9836054-A1.		
PD	20-AUG-1998.		
XX	13-FEB-1998;	98NO-AU000085.	
XX	18-NOV-1997;	97AU-0000422.	
PR	13-FEB-1997;	97AU-0005101.	
XX	(AMRA-) AMRAD OPERATIONS PTY LTD.		
PI	Antalis TM, Hooper JD;		
XX	WPI; 1998-480768/41.		
DR	P-PSDB; AAW7296.		
XX	New serine protease(s) and kinase involved in regulating cell		
PT	activity and viability - particularly the testis-specific protease		
PT	HELA2 used for modulation of fertility and as tumour suppressor		
XX	Claim 5; Fig 6; 167pp; English.		
XX	The present sequence represents the nucleotide sequence of the short		
CC	isoform of HELA2. cDNA generated from HeLa cells and PAI-2 expressing		
CC	HeLa cells was amplified using PCR primers AAV48312-13. Three new		
CC	sequences were detected in the 480 bp amplicon. These sequences are		
CC	designated HELA2 and ATC2 which have high homology to serine proteases		
CC	and BCOM3 and ATC2 which has homology to a kinase. The proteins are involved in		
CC	or associated with regulation of cell activity and/or viability.		
CC	Administration of recombinant HELA2 (also called testisin) is used to		
CC	increase fertility. Downregulation of HELA2 reduces fertility. HELA2 is		
CC	also a suppressor of testicular germ cell cancers (seminoma) and is al		
CC	expressed in some non-testicular cancers (of colon, pancreas, prostate		
CC	and ovary), so is a marker/potential therapeutic target for cancer. Th		
CC	promoter from the HELA2 gene is useful for testis-specific expression		
CC	other genes, e.g. for gene therapy or modulation of fertility. Drugs		
CC	that block activity of HELA2 should have antitumour activity (other th		
CC	in testis) while in testis recombinant HELA2 should stop growth of		
CC	tumours and normalise sperm development (eliminating the need for		
CC	orchidectomy). Identification of mutant forms of HELA2 can be used to		
CC	diagnose infertility.		
XX	Sequence 1117 BP; 241 A; 320 C; 315 G; 241 T; 0 other;		

	Query Match	97.0%	Score 1048.8;	DB 19;	Length 1117;
	Best Local Similarity	99.8%;	Pred. No. 6.6e-245;		
	Matches 1050; Conservative	0;	Mismatches 2;	Indels 0;	Gaps
QY	22	CCCGCGGAGAGAGGCCATGGCGCGCGCGGGGGCGCTGCTGCTGGCGCTGCTGCTGCCTGCTGCTGCCT	81		
Dd	1				
QY	82	CGGCTTGGACTCAGGAAGCGGAGTCGCGAGGAGCGGGCCCCATTATCAGGACCATGCGGC	141		
Dd	61	CGGCTTGGACTCAGGAAGCGGAGTCGCGAGGAGCGGGCCCCATTATCAGGACCATGCGGC	60		
QY	142	CGAGGGGTCAACGTCGCGCATCTGGTGAGAGAGCGCGCCCTTATCGGAGTGGAGGCTGCGGC	180		
Dd	121	CGAGGGGTCAACGTCGCGCATCTGGTGAGAGAGCGCGCCCTTATCGGAGTGGAGGCTGCGGC	201		
QY	202	TGGCAGGGGAGCTGGCCCTGTGGGATTTCCCACGTATCGGAGTAGAGCTGCTCAGCCAC	261		
Dd	181	TGGCAGGGGAGCTGGCCCTGTGGGATTTCCCACGTATCGGAGTAGAGCTGCTCAGCCAC	240		
QY	262	CGCTGGGCACTCACGGCGCGCACATGCTTTGAAACCTATAGTAGACCTTAGTGATCCCTCC	321		
Dd	241	CGCTGGGCACTCACGGCGCGCACATGCTTTGAAACCTATAGTAGACCTTAGTGATCCCTCC	300		
QY	322	GGGTGATGGTCAGTTTGGCCAGCTGACTTCCATGCCATGCCATCTCTGGAGCCTTGCAGGCG	381		
Dd	301	GGGTGATGGTCAGTTTGGCCAGCTGACTTCCATGCCATGCCATCTCTGGAGCCTTGCAGGCG	360		
QY	382	TACTACACCCTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCA	441		
Dd	361	TACTACACCCTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCA	420		
QY	442	CCTATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTCACTTACACTTAACACATCCAG	501		
Dd	421	CCTATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTCACTTACACTTAACACATCCAG	480		
QY	502	CCCATCTGTCTCAGGCCCTCCACAATTTAGTTTTGAGAACCAGACACTGCTGGGTGACT	561		
Dd	481	CCCATCTGTCTCAGGCCCTCCACAATTTAGTTTTGAGAACCAGACACTGCTGGGTGACT	540		
QY	562	GGCTGGGGGTACATCAAAGAGGATGGCACTGCCATCTCCCAACCCCTCCAGGAAGTT	621		
Dd	541	GGCTGGGGGTACATCAAAGAGGATGGCACTGCCATCTCCCAACCCCTCCAGGAAGTT	600		
QY	622	CAGGTCGCCATCAATAACAACCTCTATGTGCAACCACTCTTCTCAAGTACAGTTTCGC	681		
Dd	601	CAGGTCGCCATCAATAACAACCTCTATGTGCAACCACTCTTCTCAAGTACAGTTTCGC	660		
QY	682	RAGGACATCTTTGAGACATGTTTGTGTGCGAATGCCAAGGGGGGAAGATGCCTGC	741		
Dd	661	AAGGACATCTTTGAGACATGTTTGTGTGCGAATGCCAAGGGGGGAAGATGCCTGC	720		
QY	742	TTCCGTTGACTCAGGTGACCCCTTGGCCCTGTACAGAAATGGACTGTGGTATCAGATTGGA	801		
Dd	721	TTCCGTTGACTCAGGTGACCCCTTGGCCCTGTACAGAAATGGACTGTGGTATCAGATTGGA	780		
QY	802	GTCTGTAGCTGGGAGTGGGCTGTGTCGGCCCAATCGSCCGGTGTCTACACCAATATC	861		
Dd	781	GTCTGTAGCTGGGAGTGGGCTGTGTCGGCCCAATCGSCCGGTGTCTACACCAATATC	840		
QY	862	AGCCACACTTTGAGTGGATCCAGAAGCTGATGGCCCAAGAGTGGCANVTCCAGCCAGAC	921		
Dd	841	AGCCACACTTTGAGTGGATCCAGAAGCTGATGGCCCAAGAGTGGCANVTCCAGCCAGAC	900		
QY	922	CCCTCTGGCCACTACTCTTTTCCCTCTTCTGGGCTCTCCCACTCTCTGGGCGCGTC	981		
Dd	901	CCCTCTGGCGGCTACTCTTTTCCCTCTTCTGGGCTCTCCCACTCTCTGGGCGCGTC	960		
QY	982	TGAGCCTTACTGAGCCCATGCAAGCTGGGGCACTGCCAAGTCAAGGCCCTTGTCTCTTC	1041		
Dd	961	TGAGCCTTACTGAGCCCATGCAAGCTGGGGCACTGCCAAGTCAAGGCCCTTGTCTCTTC	1020		

QY	1042	TGCTTGTGGTAAATAACACATCCAGTTG	1073
Db	1021	TGCTTGTGGTAAATAACACATCCAGTTG	1052
RESULT 12			
AA	AX15337		
ID	AX15337	standard; cDNA to mRNA; 1082 BP.	
XX	AC	AA15337;	
XX	AC		
DT	04-MAY-1999	(first entry)	
XX		Nucleic acid encoding a human eosinophil serine protease.	
DE		Human; eosinophil; serine protease; allergic disease; infectious disease;	
XX		tumour; granulomatous disease; collagen disease; vascular inflammation;	
KW		ss.	
KW			
XX			
OS		Homo sapiens.	
XX			
PN	JP11032768-A.		
XX			
PD	09-FEB-1999.		
XX			
PF	16-JUL-1997; 97JP-0191319.		
XX			
PR	16-JUL-1997; 97JP-0191319.		
XX			
PA	(ONOI ) ONO PHARM CO LTD.		
XX			
DR	WPI; 1999-183825/16.		
XX			
PT		New eosinophils serine protease - useful for prevention and	
PT		treatment of allergic, infectious, tumour, granulomatous and collagen	
PT		diseases	
XX			
PS	Claim 5; Page 10-11; 18pp; Japanese.		
XX			
CC		The present sequence encodes a human eosinophil serine protease.	
CC		The protease is useful in drug compositions for the prevention and	
CC		treatment of allergic diseases, infectious diseases, tumour diseases,	
CC		granulomatous diseases, collagen diseases and vascular inflammation.	
XX			
SQ	Sequence 1082 BP; 216 A; 318 C; 308 G; 240 T; 0 other;		
Query Match 96.7%; Score 1045.8; DB 20; Length 1082;			
Best Local Similarity 99.8%; Pred. No. 3.5e-244;			
Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	31	GAGGAGGCCATGGCGCGCGCGCGCGCGTGTCTGTGGCGCTGCTGTGGCTCGGGCTGGA	90
Db	1	GAGGAGGCCATGGCGCGCGCGCGCGTGTCTGTGGCGCTGCTGTGGCTCGGGCTGGA	60
QY	91	CTCAGGAAGCCGAGTGCAGAGAGCGCGCCCTTATCAGAACCATCGGCCGACGGGTC	150
Db	61	CTCAGGAAGCCGAGTGCAGAGAGCGCGCCCTTATCAGAACCATCGGCCGACGGGTC	120
QY	151	ATCAGCTCGGCATCGTGGGTGGAGAGAGCCCACTCGGGCGTGGCGTGGCAGGGG	210
Db	121	ATCAGCTCGGCATCGTGGGTGGAGAGAGCCCACTCGGGCGTGGCGTGGCAGGGG	180
QY	211	AGCCTGCGCCCTGTGGGATTCACGCTATGCGAGTGTGCTCAGCCACCGCTGGCA	270
Db	181	AGCCTGCGCCCTGTGGGATTCACGCTATGCGAGTGTGCTCAGCCACCGCTGGCA	240
QY	271	CTCAGCGCGCGCATGCTTTGAAACCTATAGTACCTTAGTGATCCCTCCGGGTGGATG	330
Db	241	CTCAGCGCGCGCATGCTTTGAAACCTATAGTACCTTAGTGATCCCTCCGGGTGGATG	300
QY	331	GTCCAGTTTGGCCAGCTGACTTCCATGCGCATCTCTTCGAGCCTGAGGCCTACTACAC	390
Db	301	GTCCAGTTTGGCCAGCTGACTTCCATGCGCATCTCTTCGAGCCTGAGGCCTACTACAC	360

Qy 391 CGTACTTCGTATCGAATATATCTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 450  
Db 361 CGTACTTCGTATCGAATATATCTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 420  
Qy 451 ATTGCTTGGTGAAGCTGCTGCAACCTGTCACCTACCTAAACACATCCAGCCCATCTGCT 510  
Db 421 ATTGCTTGGTGAAGCTGCTGCAACCTGTCACCTACCTAAACACATCCAGCCCATCTGCT 480  
Qy 511 CTCAGGCGCTCCACATTTGAGTTCGAGAACCGGACAGACTGCTGGGTACCTGGCTGGGG 570  
Db 481 CTCAGGCGCTCCACATTTGAGTTCGAGAACCGGACAGACTGCTGGGTACCTGGCTGGGG 540  
Qy 571 TACATCAAGAGGATGAGGCACTGCCATCTCCACACCCCTCCAGGAAGTTCAGGTGCC 630  
Db 541 TACATCAAGAGGATGAGGCACTGCCATCTCCACACCCCTCCAGGAAGTTCAGGTGCC 600  
Qy 631 ATCATAAACAACCTATGTGCAACACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATC 690  
Db 601 ATCATAAACAACCTATGTGCAACACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATC 660  
Qy 691 TTTGGAGACATGGTTCGCTGGAATGCCCAAGCGGGAAGGATGCTGCTCGGTGAC 750  
Db 661 TTTGGAGACATGGTTCGCTGGAATGCCCAAGCGGGAAGGATGCTGCTCGGTGAC 720  
Qy 751 TCAGGTGGACCTTGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCTGTGAC 810  
Db 721 TCAGGTGGACCTTGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCTGTGAC 780  
Qy 811 TGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACCAATATCAGCCACCAC 870  
Db 781 TGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACCAATATCAGCCACCAC 840  
Qy 871 TTTGAGTGGATCCAGAGCTGATGCCCGCAGAGTGGCATGCCAGCCAGACCCCTCCCTGG 930  
Db 841 TTTGAGTGGATCCAGAGCTGATGCCCGCAGAGTGGCATGCCAGCCAGACCCCTCCCTGG 900  
Qy 931 CCAGTACTCTTTTCCCTCTTCTTGGGCTCTCCACTCTCCAGTCCGGGCGCGGTCTGAGCCCTAC 990  
Db 901 CCAGTACTCTTTTCCCTCTTCTTGGGCTCTCCACTCTCCAGTCCGGGCGCGGTCTGAGCCCTAC 960  
Qy 991 CTGAGCCCATCCAGCTGGGCGCACTGCCAAGTCCAGCCCTGGTCTCTTCTGCTGCTTGT 1050  
Db 961 CTGAGCCCATCCAGCTGGGCGCACTGCCAAGTCCAGCCCTGGTCTCTTCTGCTGCTTGT 1020  
Qy 1051 TGGTAATAAACACATTCAGTTGATGCT 1079  
Db 1021 TGGTAATAAACACATTCAGTTGATGCT 1049

RESULT 13  
AAD27417  
ID AAD27417 standard; DNA; 1082 BP.  
XX  
AC AAD27417;  
XX  
DT 18-APR-2002 (first entry)  
XX  
DE Human eosinophil serine protease-1 (esp-1) like enzyme DNA #2.  
XX  
KW Human; eosinophil serine protease-1; esp-1; enzyme; antiinflammatory;  
KW antihistaminic; antiallergic; osteopathic; cytostatic; dermatological;  
KW asthma; allergy; chronic obstructive pulmonary disease; COPD;  
KW osteoporosis; dermatitis; Paget's disease; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200198503-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 20-JUN-2001; 2001WO-EP06936.  
XX

21-JUN-2000; 2000US-212844P.  
31-OCT-2000; 2000US-244171P.  
30-MAR-2001; 2001US-279766P.  
(FARB ) BAYER AG.  
Xiao Y;  
WPI; 2002-122283/16.  
Novel purified human eosinophil serine protease 1-like enzyme, useful for identifying modulators of enzyme activity for treating Paget's disease, osteoporosis, airway allergy, asthma  
Disclosure; Fig 6; 131pp; English.  
The invention relates to a purified human eosinophil serine protease-1 (esp-1) like enzyme. Esp-1 like enzyme is useful in in-vitro or in-vivo assays to identify test compounds with potential therapeutic or diagnostic value. Esp-1-like enzyme modulator is useful for treating esp-1 like enzyme dysfunction related diseases condition such as asthma, airway allergy, chronic obstructive pulmonary disease (COPD) or osteoporosis. Esp-1 like enzyme is also useful in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases related to presence of mutations in the nucleic acid sequences which encode the enzyme. Pharmaceutical composition comprising esp-1 like enzyme is useful for treating dermatitis, Paget's disease, and preventing degradation of bone implants particularly dental implants. The present sequence is human esp-1 like enzyme DNA.  
Sequence 1082 BP; 216 A; 318 C; 308 G; 240 T; 0 other;  
Query Match 96.7%; Score 1045.8; DB 24; Length 1082;  
Best Local Similarity 99.8%; Pred. No. 3.5e-244;  
Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 31 GAGGAGGCCATGGGCGCGCGGGCGCTGTGCTGGCGCTGCTGCTGGCTGGCTGGCTGG 90  
Db 1 GAGGAGGCCATGGGCGCGCGGGCGCTGTGCTGGCGCTGCTGCTGGCTGGCTGGCTGG 60  
Qy 91 CTCAGGAAGCGGAGTCCGAGGAGCGCGCGCTTATCAGGACCATGCGGCGGCGGCTC 150  
Db 61 CTCAGGAAGCGGAGTCCGAGGAGCGCGCGCTTATCAGGACCATGCGGCGGCGGCTC 120  
Qy 151 ATCAGCTCGCGCATCGTGGTGGAGGAGCAGCGCAACTCGGCGCTTGGCGGCGGCGG 210  
Db 121 ATCAGCTCGCGCATCGTGGTGGAGGAGCAGCGCAACTCGGCGCTTGGCGGCGGCGG 180  
Qy 211 AGCCTGCGGCTGTGGGATTCACATGATGCGGAGTGGAGCTGCTCAGCCACCGCTGGGCA 270  
Db 181 AGCCTGCGGCTGTGGGATTCACATGATGCGGAGTGGAGCTGCTCAGCCACCGCTGGGCA 240  
Qy 271 CTCAGGCGGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCGGGTGGATG 330  
Db 241 CTCAGGCGGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCGGGTGGATG 300  
Qy 331 GTCCAGTTTGGCGCAGCTGCTTCCATGCCATCTTCTGGAGCTCGAGGCTACTTACACC 390  
Db 301 GTCCAGTTTGGCGCAGCTGCTTCCATGCCATCTTCTGGAGCTCGAGGCTACTTACACC 360  
Qy 391 CGTACTTCGTATCGAATATATCTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 450  
Db 361 CGTACTTCGTATCGAATATATCTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 420  
Qy 451 ATTGCTTGGTGAAGCTGCTGCAACCTGTCACCTACCTAAACACATCCAGCCCATCTGCT 510  
Db 421 ATTGCTTGGTGAAGCTGCTGCAACCTGTCACCTACCTAAACACATCCAGCCCATCTGCT 480  
Qy 511 CTCAGGCGCTCCACATTTGAGTTCGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGG 570  
Db 481 CTCAGGCGCTCCACATTTGAGTTCGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGG 540  
Qy 571 TACATCAAGAGGATGAGGCACTGCCATCTCCACACCCCTCCAGGAAGTTCAGGTGCC 630

Db 541 TACATCAAGAGGAGTGGGACCTCCCATCTCCACACACCTCCAGGAAGTTCAGGTGCGC 600  
QY 631 ATCAATAACAACTCTATGTGCAACCACTCTCTCTCAAGTACAGTTTCCGCAAGGACATC 690  
Db 601 ATCAATAACAACTCTATGTGCAACCACTCTCTCTCAAGTACAGTTTCCGCAAGGACATC 660  
QY 691 TTTGGAGACATGTTTGTGCTGGCAATGCCCAAGGGGGAAGATGCTCTCTCGGTGAC 750  
Db 661 TTTGGAGACATGTTTGTGCTGGCAATGCCCAAGGGGGAAGATGCTCTCTCGGTGAC 720  
QY 751 TCAGGTGACCCCTGGGCTGTACAAAGATGGAGTGTGATCAGATTGAGTGGTGGTGGC 810  
Db 721 TCAGGTGACCCCTGGGCTGTACAAAGATGGAGTGTGATCAGATTGAGTGGTGGTGGC 780  
QY 811 TGGGAGTGGGCTGTGCTGGCCCAATCGGCCGGGTCTCTACACCAATATCAGCCACCAC 870  
Db 781 TGGGAGTGGGCTGTGCTGGCCCAATCGGCCGGGTCTCTACACCAATATCAGCCACCAC 840  
QY 871 TTTGAGTGGATCCAGAGCTGATGGCCAGAGTGGGATGTCAGGACAGCCCTCTCTGG 930  
Db 841 TTTGAGTGGATCCAGAGCTGATGGCCAGAGTGGGATGTCAGGACAGCCCTCTCTGG 900  
QY 931 CCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTGCGGCGGCTGAGCCTAC 990  
Db 901 CCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTGCGGCGGCTGAGCCTAC 960  
QY 991 CTGAGCCCATGCAAGCTGGGGCCACATGCCAAGTCAGGCCCTGTTCTCTCTCTCTGTT 1050  
Db 961 CTGAGCCCATGCAAGCTGGGGCCACATGCCAAGTCAGGCCCTGTTCTCTCTCTCTGTT 1020  
QY 1051 TGGTATAAACACATTCAGTGGCT 1079  
Db 1021 TGGTATAAACACATTCAGTGGCT 1049

## RESULT 14

AAS02648

ID AAS02648 standard; cDNA; 1082 BP.

AC AAS02648;

18-JUL-2001 (first entry)

Human secreted protein gene #9.

Human secreted protein; autoimmune disorder; hyperproliferative disorder;  
cardiovascular disorder; cerebrovascular disorder; angiogenesis;  
nervous system disorder; bacterial infection; viral infection; ss;  
fungal infection; ocular disorder; wound healing; tissue regeneration;  
epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.

Homo sapiens.

WO200123547-A1.

05-APR-2001.

26-SEP-2000; 2000WO-US26337.

27-SEP-1999; 99US-0155806.

(HUMA-) HUMAN GENOME SCI INC.

Komatsoulis GA, Ruben SM, Rosen CA;

WPI; 2001-266151/27.

P-PSDB; AAU01569, AAU01595.

Nucleic acids encoding 26 human secreted polypeptides, useful for  
preventing, diagnosing and/or treating e.g. Gaucher's disease,  
Alzheimer's disease, Scimitar syndrome, Creutzfeldt-Jacob disease,  
diabetes mellitus and multiple sclerosis.

XX Sequences AAS02631-AAS02665 represent isolated nucleic acid molecules  
PS and PCR primers of the invention. Secreted proteins and their related  
XX nucleic acids can be used in the diagnosis of or susceptibility to a  
CC pathological condition by determining the presence or absence of a  
CC mutation in a nucleic acid or the presence or amount of expression of a  
CC secreted protein. The sequences are used to prevent, treat or ameliorate  
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
CC dogs, chickens or sheep. The antibodies to the polypeptides can also be  
CC used in alleviating symptoms associated with disorders and in  
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
CC immunosorbent assays (ELISA). The disorders include autoimmune diseases  
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of  
CC the breast or liver, cardiovascular disorders e.g. cardiac arrest, nervous  
CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous  
CC system disorders e.g. Alzheimer's disease, infections caused by bacteria,  
CC viruses and fungi and ocular disorders e.g. corneal infection. The  
CC peptides can also be used to aid wound healing and epithelial cell  
CC proliferation, to help prevent skin ageing due to sunburn, to maintain  
CC organs before transplantation, to regenerate tissues, in chemotaxis and  
CC as a food additive or preservative to alter storage capabilities.

SQ Sequence 1082 BP; 215 A; 318 C; 308 G; 241 T; 0 other;

Query Match 96.0%; Score 1037.8; DB 22; Length 1082;  
Best Local Similarity 99.8%; Pred. No. 3e-242;  
Matches 1039; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 39 CATGGCGCGCGGGGCGCTGCTGCTGGCGTGTGCTGGCTGGCGCTGGACTCAGGAA 98  
Db 9 CATGGCGCGCGGGGCGCTGCTGCTGGCGTGTGCTGGCTGGCGCTGGACTCAGGAA 68  
QY 99 GCGGAGTGCAGGAGGCGGCGCTTATCAGGACCATGCGGCGGAGGCTATCAGCTC 158  
Db 69 GCGGAGTGCAGGAGGCGGCGCTTATCAGGACCATGCGGCGGAGGCTATCAGCTC 128  
QY 159 GCGCATCTGCTGGTGGAGAGGAGCGCGCACTCGGGCGCTTGGCGTGGCGGAGGCTCGG 218  
Db 129 GCGCATCTGCTGGTGGAGAGGAGCGCGCACTCGGGCGCTTGGCGTGGCGGAGGCTCGG 188  
QY 219 CCGTGGGATTCACAGTATGGGAGTGGCTGTGCTAGCCACCGCTGGGCACTCAGCGC 278  
Db 189 CCGTGGGATTCACAGTATGGGAGTGGCTGTGCTAGCCACCGCTGGGCACTCAGCGC 248  
QY 279 GCGGCACTCTTGAACCTATAGTACCTAGTACCTCCCTCGGCTGGTGGTCCAGTT 338  
Db 249 GCGGCACTCTTGAACCTATAGTACCTATAGTACCTCCCTCGGCTGGTGGTCCAGTT 308  
QY 339 TGGCCAGCTGACTTCCATGCCATCTCTTGGAGCCTGAGGCTACTACACCGCTTACTT 398  
Db 309 TGGCCAGCTGACTTCCATGCCATCTCTTGGAGCCTGAGGCTACTACACCGCTTACTT 368  
QY 399 CGTATCGAATATCTATCTGAGCCCTCGTACCTGGGGAATTCACCTATGACATGCCCTT 458  
Db 369 CGTATCGAATATCTATCTGAGCCCTCGTACCTGGGGAATTCACCTATGACATGCCCTT 428  
QY 459 GGTGAAGCTGTCTGCACCTGTCCACCTACACTAAACACATCCAGCCCATCTGCTCCAGC 518  
Db 429 GGTGAAGCTGTCTGCACCTGTCCACCTACACTAAACACATCCAGCCCATCTGCTCCAGC 488  
QY 519 CTCCACATTTGAGTTTGAGAACCGGACAGTGTCTGGGTGCTGGCTGGGGGTACATCAA 578  
Db 489 CTCCACATTTGAGTTTGAGAACCGGACAGTGTCTGGGTGCTGGCTGGGGGTACATCAA 548  
QY 579 AGAGGATGAGGACTGCCATCTCCACACCGCTCCAGGAAGTTCAGGTGCGCCATCAAAA 638  
Db 549 AGAGGATGAGGACTGCCATCTCCACACCGCTCCAGGAAGTTCAGGTGCGCCATCAAAA 608  
QY 639 CAACTCTATGTGCAACCACTCTTCTCAAGTACAGTTTTCGCAAGGACATCTTTGGAGA 698  
Db 609 CAACTCTATGTGCAACCACTCTTCTCAAGTACAGTTTTCGCAAGGACATCTTTGGAGA 668

QY	699	CATGGTTTGCTGGCAATGCCCAAGCGGGGAAGGATGCCTTCGCTGACTCAGGTGG	758
Db	669	CATGGTTTGCTGGCAATGCCCAAGCGGGGAAGGATGCCTTCGCTGACTCAGGTGG	728
QY	759	ACCTTTGGCCTGTAAACAAGATGGACTGTGGTATCAGATTTGGAGTCGTGAGCTGGGGAGT	818
Db	729	ACCTTTGGCCTGTAAACAAGATGGACTGTGGTATCAGATTTGGAGTCGTGAGCTGGGGAGT	788
QY	819	GGGCTGTGTGGCGCCCAATCGGCCGGGTCTTACACCAATATCAGCCACCACCTTTGAGTG	878
Db	789	GGGCTGTGTGGCGCCCAATCGGCCGGGTCTTACACCAATATCAGCCACCACCTTTGAGTG	848
QY	879	GATCCAGAAGCTGATGGCCAGAGTGGCATGTCCTCAGCCAGACCCCTCTCTGGCCACTACT	938
Db	849	GATCCAGAAGCTGATGGCCAGAGTGGCATGTCCTCAGCCAGACCCCTCTCTGGCCACTACT	908
QY	939	CTTTTTCCTCTCTCTGGGCTCTCCCACTCTCTGGGCGCGGTCTGAGCCTACCTGAGCCC	998
Db	909	CTTTTTCCTCTCTCTGGGCTCTCCCACTCTCTGGGCGCGGTCTGAGCCTACCTGAGCCC	968
QY	999	ATCAGCCTGGGCCCACTGCCAAGTCAGGCCCTGGTTCTCTCTCTTTTGGTAATA	1058
Db	969	ATCAGCCTGGGCCCACTGCCAAGTCAGGCCCTGGTTCTCTCTCTTTTGGTAATA	1028
QY	1059	AACACATTCAGTTGATGCCT	1079
Db	1029	AACACATTCAGTTGATGCCT	1049
RESULT 15			
AA15336			
ID	AA15336 standard; cDNA to mRNA; 942 BP.		
XX	XX		
AC	AA15336;		
XX	XX		
DT	04-MAY-1999 (first entry)		
XX	XX		
DE	Nucleic acid encoding a human eosinophil serine protease.		
XX	XX		
KW	Human; eosinophil; serine protease; allergic disease; infectious disease		
KW	tumour; granulomatous disease; collagen disease; vascular inflammation		
XX	SS.		
OS	Homo sapiens.		
XX	XX		
Key	Location/Qualifiers		
FT	1..942		
CDS	/*tag- a		
FT	/note= "no termination codon"		
FT	JP11032768-A.		
PN	XX		
XX	XX		
PD	09-FEB-1999.		
XX	XX		
PF	16-JUL-1997; 97JP-0191319.		
XX	XX		
PR	16-JUL-1997; 97JP-0191319.		
XX	XX		
PA	(ONOY ) ONO PHARM CO LTD.		
XX	XX		
DR	WPI; 1999-183825/16.		
DR	P-PSDB; AAW97116.		
XX	XX		
PT	New eosinophils serine protease - useful for prevention and		
PT	treatment of allergic, infectious, tumour, granulomatous and collagen		
PT	diseases		
XX	XX		
PS	Claim 4; Page 10; 18pp; Japanese.		
XX	XX		
CC	The present sequence encodes a human eosinophil serine protease.		
CC	The protease is useful in drug compositions for the prevention and		
CC	treatment of allergic diseases, infectious diseases, tumour diseases,		

CC	granulomatous diseases, collagen diseases and vascular inflammation.
XX	
50	Sequence 942 BP; 176 A; 283 C; 276 G; 207 T; 0 other;
	Query Match 86.8%; Score 938.8; DB 20; Length 942;
	Best Local Similarity 99.8%; Pred. No. 3.1e-218;
	Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps
Qy	40 ATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGGCTGACTCAGGAAG 99
Db	1 ATGGCGCGCGGGCGGCTGCTGCTGGCGCTGCTGCTGGCTGGGCTGGACTCAGGAAG 60
Qy	100 CCGAGTGCAGGAGGCGGCGCCCTTATCAGGACCATCGCGCGCGACGGGTCATCAGCTCG 159
Db	61 CCGAGTGCAGGAGGCGGCGCGCTTATCAGGACCATCGCGCGCGACGGGTATCATCAGCTCG 120
Qy	160 CGCATCGTGGGTGGAGGAGCGCGCAACTCGGGCGTTGGCCGTGGCGAGGGAGCGCTCGCG 219
Db	121 CGCATCGTGGGTGGAGGAGCGCGCAACTCGGGCGTTGGCCGTGGCGAGGGAGCGCTCGCG 180
Qy	220 CTGTGGATTCGCCAGCTATCGGGAGTCAGCCTGCTCAGCCACCGCTGGGCACTCACGGCG 279
Db	181 CTGTGGATTCGCCAGCTATCGGGAGTCAGCCTGCTCAGCCACCGCTGGGCACTCACGGCG 240
Qy	280 GCGCAGCTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTT 339
Db	241 GCGCAGCTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTT 300
Qy	340 GGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGAGGCCCTACTACACCGCTTACTTC 399
Db	301 GGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCGAGGCCCTACTACACCGCTTACTTC 360
Qy	400 GTATCGAATATCTATCTGAGCCCTCGCTACTCTGGGGAATTCACCCCTATGACATTCGCTTG 459
Db	361 GTATCGAATATCTATCTGAGCCCTCGCTACTCTGGGGAATTCACCCCTATGACATTCGCTTG 420
Qy	460 GTGAAGCTGTCTGCACCTGTCACTAACAACACATCCAGGCCCATCTGCTTCCAGGCC 519
Db	421 GTGAAGCTGTCTGCACCTGTCACTAACAACACATCCAGGCCCATCTGCTTCCAGGCC 480
Qy	520 TCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA 579
Db	481 TCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA 540
Qy	580 GAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGTGCGCATCATAAAC 639
Db	541 GAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGTGCGCATCATAAAC 600
Qy	640 AACTCTATGTGCACACCACTTCTCCTCAAGTACAGTTTCCGAAGGACATCTTTGGAGAC 699
Db	601 AACTCTATGTGCACACCACTTCTCCTCAAGTACAGTTTCCGAAGGACATCTTTGGAGAC 660
Qy	700 ATGTTTGTGTGTGCAATGCCAAGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGA 759
Db	661 ATGTTTGTGTGTGCAATGCCAAGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGA 720
Qy	760 CCCTTGGCGCTGTAAACAGATGGAAGCTGTGTGTATCAGATTCGAGCTCGGGAGTG 819
Db	721 CCCTTGGCGCTGTAAACAGATGGAAGCTGTGTGTATCAGATTCGAGCTCGGGAGTG 780
Qy	820 GGCTGTGTGGGCCCAATCGGCCCGGTGCTACACCAATATACGCCACCACTTTTGAGTGG 879
Db	781 GGCTGTGTGGGCCCAATCGGCCCGGTGCTACACCAATATACGCCACCACTTTTGAGTGG 840
Qy	880 ATCCAGAAGCTGTATGGCCCAAGTGGCATGTCCAGCCAGACACCCCTCTGGGCCACTACTC 939
Db	841 ATCCAGAAGCTGTATGGCCCAAGTGGCATGTCCAGCCAGACACCCCTCTGGGCCACTACTC 900
Qy	940 TTTTTCCTCTCTCTGGGCTCTCCCACTCCTGGGGCCGGTC 981
Db	901 TTTTTCCTCTCTCTGGGCTCTCCCACTCCTGGGGCCGGTC 942

Mon May 12 16:37:38 2003

Search completed: May 10, 2003, 03:24:32  
Job time : 355 secs

us-09-787-844-1.rng

---

Page 16



STIC-Biotech/ChemLib

93059

From: Yu, Misook  
Sent: Friday, May 02, 2003 1:50 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/787,844

Please search SEQ ID NO:1.

Examiner Misook Yu, Ph.D.  
703-308-2454 (Phone)  
Art Unit 1642  
CM1-8E18 (Room)  
CM1-8E12 (Mail Box)

Mary Jane Ruhl  
Tech. Info. Specialist, STIC  
TC-1600  
CM-1, Room 6A-06  
Phone: 605-1155

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 5/2/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 03:17:32 ; Search time 75 Seconds  
(without alignments)  
4420.234 Million cell updates/sec

Title: US-09-787-844-1

Perfect score: 1081

Sequence: 1 agaggcagagggggcgctca.....acattccagtgatgctgc 1081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_NA.\*
- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
  - 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
  - 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
  - 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
  - 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	100.0	1081	4	US-09-008-271A-15
2	752.6	69.6	1103	4	US-09-386-642-59
3	165	15.3	1110	4	US-09-386-653A-1
4	155.8	14.4	1130	4	US-09-386-653A-8
5	131.4	12.2	1430	4	US-09-386-629-1
6	126.8	11.7	1166	4	US-09-386-629-2
7	118.6	11.0	1108	2	US-09-016-366A-14
8	118.6	11.0	1108	2	US-08-978-404B-20
9	117.8	10.9	1142	4	US-09-386-642-8
10	117.8	10.9	1169	4	US-09-386-642-7
11	117	10.8	1137	2	US-09-016-366A-18
12	117	10.8	1137	2	US-08-978-404B-13
13	116.2	10.7	1783	4	US-09-510-738A-188
14	116.2	10.7	2363	4	US-09-742-703-3
15	115.4	10.7	735	4	US-09-079-970A-1
16	115.4	10.7	771	4	US-09-079-970A-4
17	115.4	10.7	1081	2	US-09-016-366A-22
18	115.4	10.7	1081	2	US-08-978-404B-17
19	115.4	10.7	1128	2	US-09-016-366A-20
20	115.4	10.7	1128	2	US-08-978-404B-15
21	115.4	10.7	1154	2	US-09-016-366A-16
22	115.4	10.7	1154	2	US-08-978-404B-11
23	115.4	10.7	1219	2	US-08-978-404B-7
24	112	10.4	2413	4	US-09-518-046-1
25	111.2	10.3	1605	2	US-09-000-846-1
26	106	9.8	1031	2	US-08-978-404B-1
27	105	9.7	654	4	US-09-280-116-57

28	104.2	9.6	1097	2	US-08-978-404B-4
29	104	9.6	1155	4	US-09-163-951-15
30	104	9.6	1240	4	US-09-163-951-14
31	102.6	9.5	1103	2	US-09-016-366A-24
32	100	9.3	1109	4	US-09-088-651-6
33	97	9.0	1109	4	US-09-088-651-1
34	96	8.9	2416	4	US-09-261-416-1
35	95.8	8.9	683	4	US-09-088-651-3
36	93.8	8.7	1095	2	US-08-978-404B-9
37	93.8	8.7	2544	4	US-09-518-046-3
38	85.6	7.9	1440	1	US-07-882-202A-3
39	85.6	7.9	1440	1	US-08-021-615A-3
40	85.6	7.9	1440	1	US-08-321-777-3
41	85.6	7.9	1440	3	US-09-009-217-13
42	85.6	7.9	1440	3	US-09-009-656-13
43	85.6	7.9	1440	5	PCT-US93-04493-3
44	85.6	7.9	2422	1	US-08-475-845-1
45	85.6	7.9	2422	1	US-08-327-690-1

ALIGNMENTS

RESULT 1  
US-09-008-271A-15  
Sequence 15, Appli  
Patent No. 6203979  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
Hillman, Jennifer L.  
Yue, Henry  
Guegler, Karl J.  
Corley, Neil C.  
Tang, Tom Y.  
Shan, Purvi  
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,271A  
FILING DATE: 16-Jan-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Mohan-Peterson, Sheela  
REGISTRATION NUMBER: 41,201  
REFERENCE/DOCKET NUMBER: PF-0458 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1081 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSTAT03  
CLONE: 789927  
SEQUENCE DESCRIPTION: SEQ ID NO: 15 :  
US-09-008-271A-15

Query Match									
Best Local Similarity 100.0%; Score 1081; DB 4; Length 1081;									
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps									
QY	1	AGAGG	CAGAGGGGGCGT	CAGGCCGCGGGAGAGGAGGCAT	TGGCGCGCGGGGCGCGCTG	60			
DB	1	AGAGG	CAGAGGGGGCGT	CAGGCCGCGGGAGAGGAGGCAT	TGGCGCGCGGGGCGCGCTG	60			
QY	61	CTGCT	CGCGCTCGT	CGCTCGGCTGGACTCAGGAAGCCGGAGTCGAGAGGCGGCG	120				
DB	61	CTGCT	CGCGCTCGT	CGCTCGGCTGGACTCAGGAAGCCGGAGTCGAGAGGCGGCG	120				
QY	121	CCCTT	ATCAGGACCATGCGCGCGACGGGTCACT	CAGCTCGGCATCTGCTGGGTGGAGAGGAC	180				
DB	121	CCCTT	ATCAGGACCATGCGCGCGACGGGTCACT	CAGCTCGGCATCTGCTGGGTGGAGAGGAC	180				
QY	181	GCGGA	ACTCGGGCGTTGGCGGTGGCAGGGAGCCCTGCGCCTTGGGATATCCACGTATGC	240					
DB	181	GCGGA	ACTCGGGCGTTGGCGGTGGCAGGGAGCCCTGCGCCTTGGGATATCCACGTATGC	240					
QY	241	GGAGT	GAGCGCTCTCAGCACCGCTGGGCACCTACACGGGGGCGACCTGCTTTGAAACCTAT	300					
DB	241	GGAGT	GAGCGCTCTCAGCACCGCTGGGCACCTACACGGGGGCGACCTGCTTTGAAACCTAT	300					
QY	301	AGTGAC	CTTAGTGATCCCTCCGGGTGGATGGTGCAGTTTGGCAGCTGACTTCCCATGCCA	360					
DB	301	AGTGAC	CTTAGTGATCCCTCCGGGTGGATGGTGCAGTTTGGCAGCTGACTTCCCATGCCA	360					
QY	361	TCCCT	CTGGAGCCTGCAGGCCCTACTACACCGGTACTTCGATATCGAATATCTATCTGAGC	420					
DB	361	TCCCT	CTGGAGCCTGCAGGCCCTACTACACCGGTACTTCGATATCGAATATCTATCTGAGC	420					
QY	421	CCTGC	CTACCTGGGGAATTACCCCTATGACATTTGCCTTGGTGAAGCTGCTGCACCTGTC	480					
DB	421	CCTGC	CTACCTGGGGAATTACCCCTATGACATTTGCCTTGGTGAAGCTGCTGCACCTGTC	480					
QY	481	ACCTA	CATAAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTTGCAGTTTGGAGAC	540					
DB	481	ACCTA	CATAAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTTGCAGTTTGGAGAC	540					
QY	541	CGGAC	AGACTGCTGGGTGACTGGCTGGGGTACATCAAGAGGATGAGGCATGCGCATCT	600					
DB	541	CGGAC	AGACTGCTGGGTGACTGGCTGGGGTACATCAAGAGGATGAGGCATGCGCATCT	600					
QY	601	CCCCA	CACTCCAGGAAGTTTCAGGTGCGCATCAATAACACTCTATGTGGAACACCTC	660					
DB	601	CCCCA	CACTCCAGGAAGTTTCAGGTGCGCATCAATAACACTCTATGTGGAACACCTC	660					
QY	661	TTCTC	TAAGTACAGTTTCGCCAAGGACATCTTTGGAGACATGGTTGTGCTGGCAATGCC	720					
DB	661	TTCTC	TAAGTACAGTTTCGCCAAGGACATCTTTGGAGACATGGTTGTGCTGGCAATGCC	720					
QY	721	CAAGG	CGGGAAGCATGCGTTCGGTGACTCAGGTGGACCCCTTGGCCTGTAAACAAGAT	780					
DB	721	CAAGG	CGGGAAGCATGCGTTCGGTGACTCAGGTGGACCCCTTGGCCTGTAAACAAGAT	780					
QY	781	GGACT	GTGTATCAGATTGGAGTCGTGAGCTGGGAGTGGGCTGTGGTGGGCCCAATCGG	840					
DB	781	GGACT	GTGTATCAGATTGGAGTCGTGAGCTGGGAGTGGGCTGTGGTGGGCCCAATCGG	840					
QY	841	CCGGT	GTCTACACCAATATCAGCCACCACCTTTGAGTGGATCCAGAAGCTGATGCCCCAG	900					
DB	841	CCGGT	GTCTACACCAATATCAGCCACCACCTTTGAGTGGATCCAGAAGCTGATGCCCCAG	900					
QY	901	AGTGG	CACTGCCAGACCCCTCTGGGCCCATCTCTTTTTCCCTCTTCTCTGGGCT	960					
DB	901	AGTGG	CACTGCCAGACCCCTCTGGGCCCATCTCTTTTTCCCTCTTCTCTGGGCT	960					
QY	961	CTCCA	CTCTGGGCGGCTGTGAGCCTACCTGAGCCCATCGACGCTGGGGCACA	1020					
DB	961	CTCCA	CTCTGGGCGGCTGTGAGCCTACCTGAGCCCATCGACGCTGGGGCACA	1020					

```

Qy 1021 ACTCAGGCCCTGGTCTCTTCTGCTCTGTGGTAAATAAACACATTCACGTTGATGCGCTG 1080
      |||||
Db 1021 AGTCAGGCCCTGGTCTCTTCTGCTCTGTGGTAAATAAACACATTCACGTTGATGCGCTG 1080
      |||||
Qy 1081 C 1081
      |
Db 1081 C 1081

RESULT 2
US-09-386-642-59
; Sequence 59, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jenson
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
; OTHER INFORMATION: sequence of human protease F in CFEX2 zymogen
; OTHER INFORMATION: vector
US-09-386-642-59

```

Query Match	69.6%;	Score 752.6;	DB 4;	Length 1103;
Best Local Similarity	93.6%;	Pred. No. 5,1e-186;		
Matches 785; Conservative	0;	Mismatches 54;	Indels	Gaps
QY 163	ATCGTGGGTGAGAGACGCCAACTCGGGCGTTGGCGTGCCAGGGAGCCTGC GCCTG	222		
Db				
QY 139	ATCGTTGGGGCTATGCTTAGAATCTGGGGCGTTGGCGTGCCAGGGAGCCTGC GCCTG	198		
Db				
QY 223	TGGGATTCCCACGTATGGGGAGTGAGCCTGTCTCAGCCACCCTGGTGGCACTCACGSCGCG	282		
Db				
QY 199	TGGGATTCCACGTTATGCGGAGTGAGCCTGCTCACCAACCGCTGGCACTCACGGCGCG	258		
Db				
QY 283	CACCTGCTTTGAACCATATAGTGACCTTAGTGATCCCTCCGGTGGAGTTCAGTTTGGC	342		
Db				
QY 259	CACCTGCTTTGAACCATATAGTGACCTTAGTGATCCCTCCGGTGGAGTTCAGTTTGGC	318		
Db				
QY 343	CAGCTGACTTCCATGCCATCTCTCTGGAGCCTGCAGGCCTACTACACCGTTACTTCGTA	402		
Db				
QY 319	CAGCTGACTTCCATGCCATCTCTCTGGAGCCTGCAGGCCTACTACAACCGTTACTTCGTA	378		
Db				
QY 403	TCGAATAFCTATCTGAGCCCTCGCTACCTGCGGGAATTCACCTATGACATTCGCTTGGTG	462		
Db				
QY 379	TCGAATAFCTATCTGAGCCCTCGCTACCTGCGGGAATTCACCTATGACATTCGCTTGGTG	438		
Db				
QY 463	AAGCTGTCTGCACCTGTCACTTACACTAAACACATCAGCCCATCTCTCTCCAGGCCTCC	522		
Db				
QY 439	AAGCTGTCTGCACCTGTCACTTACACTAAACACATCAGCCCATCTCTCTCCAGGCCTCC	498		
Db				
QY 523	ACATTTGAGTTTGAGAACC GGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAAAGAG	582		
Db				
QY 499	ACATTTGAGTTTGAGAACC GGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAAAGAG	558		
Db				
QY 583	GATGAGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGGTGGCCATCATAAAAC	642		
Db				
QY 559	GATGAGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGGTGGCCATCATAAAAC	618		
Db				
QY 643	TCATATGCAACCACTTCCTCAAGTACAGTTTCGGGAAGACATCTTTGGAGACATG	702		
Db				
QY 619	TCATATGCAACCACTTCCTCAAGTACAGTTTCGGGAAGACATCTTTGGAGACATG	678		
Db				

```
QY 703 GTTTGTGCTGCGCAATGCCAAGCGGGAAGATGCCCTGCTCGTGTGACTCAGGTGGACCC 762
Db 679 GTTTGTGCTGCGCAATGCCAAGCGGGAAGATGCCCTGCTCGTGTGACTCAGGTGGACCC 738
QY 763 TTGGCCCTGTACAGAGATGACTGTGTATCAGATTGGAGTCTGAGCTGGGAGTGGCC 822
Db 739 TTGGCCCTGTACAGAGATGACTGTGTATCAGATTGGAGTCTGAGCTGGGAGTGGCC 798
QY 823 TGTGCTGCGGCCAATCGCGCCGCTGTCTACACCAATATCAGCCACCACTTTGAGTGGATC 882
Db 799 TGTGCTGCGGCCAATCGCGCCGCTGTCTACACCAATATCAGCCACCACTTTGAGTGGATC 858
QY 883 CAGAGCTGATGGCCAGAGTGGCATGTCCAGCCAGCCCTCTCTGCCCACTACTCTTT 942
Db 859 CAGAGCTGATGGCCAGAGTGGCATGTCCAGCCAGCCCTCTCTGCCCACTACTCTTT 918
QY 943 TTCCCTCTCTCTGCTGCTCCCACTCTCTGGGCGGCTCTGAGCTACCTGAGCCCATG 1001
Db 919 CATCACCATCACTAGCGCGCGCTTCCCTTTAGTGAGGGTTAAATGCTTCGAGCAGACATG 977

RESULT 3
US-09-386-653A-1
; Sequence 1, Application US/09386653A
; Patent No. 6458564
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Barrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/09/386, 653A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-386-653A-1

Query Match 15.3%; Score 165; DB 4; Length 1110;
Best Local Similarity 53.7%; Pred. No. 1.1e-33;
Matches 429; Conservative 0; Mismatches 340; Indels 30; Gaps 3;

QY 113 AGCGCGCCCTTATCAGGACCATGCGCGGAGCGGTATCATCGTCCGCGATCGTGGTG 172
Db 88 AGAGGGCCAAAGGACCAACAGCCCTGTGCTGCCCGCCAGGATGCTGAACCGAATGTTGGCG 147
QY 173 GAGAGACGCGCAACTCGGCGTGTGGCGTGGCAGGAGGAGCTCGCGCTGTGGGATTCCT 232
Db 148 GCGAGACACGCGAGGCGGAGTGGCCCTGGCAAGTACAGATCCAGCCAGCGCAAGGAGCC 207
QY 233 ACGTATGCGAGTGGAGCTGCTCAGCAGCCGCTGGGCACTCAGCGCGCGCACTGCTTTG 292
Db 208 ACTTCTCGGGGCGAGCTCATCGCGAGCAGTGGTCTCTGACGCTGCGCACTGCTTCC 267
QY 293 AAACCTATAGTACCTTAGTGTATCCCTCCGCGTGGATGCTCAGTTGGCCAGCTGACTT 352
Db 268 GCAACACCTCTGAGC-----GTCCCTGTACCAAGTCTCTGCTGGGGCAA 312
QY 353 CCATGCCATCTTCTGAGGCGCTGAGGCGCTACTACACCCGCTTACTTCGTATCGAATATCT 412
Db 313 GCGAGCTAGTGCAGCGGAGCACACCGCTATGTATGCCCGGGTGGAGCAGGTGGAGACA 372
QY 413 ATCTGAGCCCTCGCTACCTGGGAATTCACCCCTATGACATTCCTTGGTGAAGTCTGCTG 472
Db 373 ACC---CCCTGTACCGGCGAGCGCTCCAGCGCTGAGCTGGCCCTGCTGGAGCTGGAGG 429
QY 473 CACCTGTACCTACATAACACATCCAGGCCATCTCTCTCCAGGCGCTCCACATTTAGT 532
Db 430 CACAGTGCCTTACACCAATATACATCTCCCGCTGTGCCTGCCTGACCCCTCGGTGATCT 489
```

```
QY 533 TTGAGAACCGACAGACTGCTGGTGTGCTGGGGTACATCAAGAGGATGAGGCAC 592
Db 490 TTGAGACGGGCATCAACTGCTGGTGTGCTGGGGTACATCAAGAGGATGAGGCAC 549
QY 593 TGGCATCTCCCAACACCTCCAGGAAGTTCCAGTTCGCGCATCAATAAACAATCTATGTGCA 652
Db 550 TGGCCGAACCGCGGATCTCGAGAACTCGCTGTGCCCATCATCGACACACCAAGTGCA 609
QY 653 ACCACCTCTTCCTCAAGTACA-----GTTTCCGCAAGGACATCTTTGGAGACA 700
Db 610 ACCTGCTCTACAGAAAGACACCGAGTTGGCTACCAACCCCAAAACCATCAAGATGACA 669
QY 701 TGGTTTGTGCTGCAATGCCCAAGCGGGAAGTATGCTTCTCGTGTGCTGACTCAGGTGGAC 760
Db 670 TGTGTGCGCGGCTTCGAGGAGGCAAGAGGATGCTGCAAGGGCGACTCGGGCGGCC 729
QY 761 CTTTGGCTGTAAACAAGATGGAGTGTGTTATCAGATTGGAGTCTGAGCTGGGGAGTGG 820
Db 730 CCCTGGTGTGCTGCTGGGTGAGTGTGCTGAGGCGGGGTGATCAGCTGGGGTGAAG 789
QY 821 GCTGTGTGCGCCCAATCGCGCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGGA 880
Db 790 GCTGTGCGCGCCAGAACCGCCAGGTGTCTACATCCGTGTACCGCCCAACCAACTGGGA 849
QY 881 TCCAGAAGCTGATGGCCCA 899
Db 850 TCCATCGGATCATCCCA 868

RESULT 4
US-09-386-653A-8
; Sequence 8, Application US/09386653A
; Patent No. 6458564
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Barrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/09/386, 653A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of
; OTHER INFORMATION: Protease T in a zymogen activation vector
US-09-386-653A-8

Query Match 14.4%; Score 155.8; DB 4; Length 1130;
Best Local Similarity 60.4%; Pred. No. 2.6e-31;
Matches 281; Conservative 0; Mismatches 172; Indels 12; Gaps 1;
```

```
QY 447 TGACATTCCTGTTGGTGAAGTGTGTCACCTGTGTCACCTACACTAAACATCCAGCCAT 506
Db 432 TGAGTGGCCCTGTTGGAGCTGGAGCCAGCCAGTGCCTTCCACCAATTACATCTCCCGT 491
QY 507 CTGTCTCCAGGCTCCACATTTGAGTTGAGAACCGGACAGACGCTGGGTGACTGGCTG 566
Db 492 GTGCTCTGCTGACCCCTCGGTGATCTTTGAGACGGGATGAACTGCTGGGTCACTGGCTG 551
QY 567 GGGGTACATCAAGAGGATGAGGCATCTCCCATCTCCCAACACCTCCAGGAAGTTCAGT 626
Db 552 GGGCAGCCCGAGTGAAGAGACCTCTTCCCGAACCCGGATCTCTGCAGAACTCGCTGT 611
QY 627 CGGCATCAATAAACAATCTATGTGCAACCACTCTTCTCTCAAGTACA-----G 674
Db 612 GCCCATCATGACACACCCCAAGTCAACCTGCTCTACAGCAAGACACCGAGTTGGCTA 671
```

Query Match	11.7%	Score 126.8;	DB 4;	Length 1166;
Best Local Similarity	49.6%;	Pred. No. 8.6e-24;		
Matches 363;	Conservative 0;	Mismatches 357;	Indels 12;	Gaps 1;

  

QY	163	ATCGTGGGTGGAGGACGCCGAAC	TCGGCGCTTGGCGCTGGCAGGGGACGCTGCGCGCTG	222
Db	166	ATCGTTGGGGCTATGCTCT	FAGAGGACAGCAGTGGCCCTGGATCTGAGCATCCGAAG	225
QY	223	TGGGATTCGCCACGTATCGGAGT	GAGCCTGCTCAGCCACCGCTGGGCACCTACAGGGGGCG	282
Db	226	AATGGACCCACACACTCGCAGG	TTCTCTCTCACCAGCGCTGGTGATCATTGCTGCC	285
QY	283	CAGTCGCTTTGAACCTPATAGTG	ACCTTATGTATCCCTCGGGGTGGATGTTCCAGTTTGGC	342
Db	286	CAGCTGTTTCAAGGACAACTG	AACAACCATACCTGTTCTCTGTGCTGGGGGCGCTGG	345
QY	343	CAGCTGACTTCCATGCCATCCT	CTCTGGAGCCTGCAGGCGCTACTACACCCGTACTTCGTGA	402
Db	346	CAGCTGGGGAACCTTGGCTCT	-----CGGTCCAGAGTGGGTGTGGCCTGG	393
QY	403	TCGATATCTATCTAGCGCCTCG	CTACCTGCGGGAATTCACCCATGACATTGCCTTGGTG	462
Db	394	GTGGAGCCGCCACCTGTGTAT	TCTCTGGAAGGAGTGCCTGTGTGCAGACATTGCCCTGGTG	453
QY	463	AAGCTGTCTGCAACCTGTCACT	TACACTAAACACATCCAGGCCCATCTGTCTCCAGGCCCTCC	522

Db 454 CGTCTCGAGCGCTCCATACAGTTCTCAGACGGGCTCTGCCCATCTGCTACCTGATGCC 513  
QY 523 ACATTTGAGTTTGAGAACCGGACAGCTGCTGGGTGACTGGCTGGGGGTACATCAAGAG 582  
Db 514 TCTATCCACCTCCCTCCAAACACCCACCTGCTGGATCTCAGGCTGGGGAGCATCCAGAT 573  
QY 583 GATGAGGACATGCCATCTCCACACACCCTCCAGGAAGTTCAGGTGCGCCATCATAAACAC 642  
Db 574 GGAGTTCCCTTGGCCCACTCCAGACCTCCAGACCTCCAGAAAGCTGAAGTTCTATCATCGACTCG 633  
QY 643 TCTATGTCGACACCACTCTTCTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG 702  
Db 634 GAAGTCTCCACCCATCTCTACTGCGGGGAGCAGGACGAGGACCCATCACTGAGGACATG 693  
QY 703 GTTCTGCTGGCAATGCCAAGCGGGGAGGATGCTCTCTGGGTGACTCAGGTGACCC 762  
Db 694 CTGTGTCGGGTACTTGGAGGGGAGGGGATGCTTCTGGGGGAGCTCCGGGGGCCCC 753  
QY 763 TTGGCCTCTAAACAAGATGGACTGTGTGATCAGATTGGAGTCTGAGCTGGGGAGTGGGC 822  
Db 754 CTCATGTGCCAGGTGGACGGCGCTGCTGCTGGCGGCATCATCAGCTGGGGCGAGGGC 813  
QY 823 TGTGTCGGCCCACTCGCCCGGTCTCTACACCAATATCAGCCACCACTTTGAGTGGATC 882  
Db 814 TGTGCGGAGCGCAACAGCCCGGGGTCTACATCAGCCTCTCTGCGCACCGCTCTCTGGGTG 873  
QY 883 CAGAACTGATG 894  
Db 874 GAGAGATCGTG 885

RESULT 7

US-09-016-366A-14  
; Sequence 14, Application US/09016366A  
; Patent No. 5955431  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,366A  
; FILING DATE: January 30, 1998  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/037,090  
; FILING DATE: 05-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7093  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1108 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-016-366A-14  
Query Match 11.0%; Score 118.6; DB 2; Length 1108;  
Best Local Similarity 56.0%; Pred. NO. 1.1e-21;  
Matches 248; Conservative 0; Mismatches 189; Indels 6; Gaps 1;  
QY 448 GACATTGGCTTGGTGAAGCTGTCTGCACCTGTCACTACATTAACACATCCAGCCCATC 507  
Db 397 GAGGTTCCTGCTGGAGCTGTGAGTCCCTGTGAATGTCTCCACCATATCCACCCCATATA 456  
QY 508 TGTCTCCAGGCTCCACATTTTTCAGTTTGAAGAACCGGACAGACTGTGGGTGACTGGCTGG 567  
Db 457 TCCCTGCCCCCTCGCTCGGAGACCTTCCCTCCCTGGACATCTGCTGGGTGACAGCTGG 516  
QY 568 GGTGATACAAAGAGATGAGGCACTGCCATCTCCACACACCTCCAGGAAGTTCAGGTTC 627  
Db 517 GCGGACATTAATAATGACGAGCCTCTCCACCTCTCTATCTCTCTGAAGCAAGTGAAGTT 576  
QY 628 GCCATCATAAACAACACTCTATGTGCAACCACTCTTCTCAAGTACAGTTTCCGCAAGGAC 687  
Db 577 CCCATGTGGAACACAGCCTGTGTGACCGGAAGTACACACTGGGCTCTACACGGAGAT 636  
QY 688 ATCTTTGGAGACATGGTTTGTG-----CTGGCAATGCCCAAGCGGGAGGATCCCTGC 741  
Db 637 GATTTTCCCATTTGCTCATGATGGCATGCTGTGTGCTGGAAATACACGAGAGAGACTCCTGC 696  
QY 742 TTCGGTGACTCAGGTGGACCCCTTGGCCCTGTACCAAGATGAGTGTGGTATCAGATTGGA 801  
Db 697 CAGGGCGATTTCAGGGGGCCACTGCTGTCTGCAAAAGTGAAGGGTACCTGGCTGACGACGGA 756  
QY 802 GTCGTGAGCTGGGGAGTGGCTGTGGTGGCCCAATCGCCCGGTGTCTACACCAATATC 861  
Db 757 GTGTGACCTGGGGTGGAGGCTGCGCACAGCCCAACAAAGCTTGGCATCTACACCCGGGTG 816  
QY 862 AGCCACCACTTTGAGTGGATCCA 884  
Db 817 ACATACTACTTAGACTGGATCCA 839

RESULT 8

US-08-978-404B-20  
; Sequence 20, Application US/08978404B  
; Patent No. 5968782  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
; TITLE OF INVENTION: FIBRINOGEN  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,404B  
; FILING DATE: 25-NOV-97  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/032,354  
; FILING DATE: 04-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7090



```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
;
; TEXT:
;
; INFORMATION FOR SEQ ID NO: 20:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-978-404B-20

```

Query Match 11.0%; Score 118.6; DB 2; Length 1108;  
Best Local Similarity 56.0%;  
Pred. No. 1.1e-21;  
Matches 248; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

448	QY	GACATTTGGCTGGTGAAGCTGTCTGCACCTGTCCACCTACACTAAACACATCCAGCCCATC	507
397	Db	GACGTTTGGCTGTCTGGAGCTTGAGGTCCTGTCAATCTCCACCATATCCACCCCATTA	456
508	QY	TGTCCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGG	567
457	Db	TCCCTGCCCTTGCTTCGGAGACCTTCCCCCCTGGGACATCTGCTGGGTGCACAGCTGG	516
568	QY	GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGTTC	627
517	Db	GGCGACATGATTAATGACGAGCGCTCTCCCACTCTTATCTCTGAAGCAAGTGAAGTT	576
628	QY	GCCATCATAAACAACCTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCAAGCAC	687
577	Db	CCATTTGGGAAAACAGCGCTGTGTGACGGGAAGTACCACTGGCTCTACACGGGAGAT	636
688	QY	ATCTTTGGAGACATGGTTTGTG-----CTGCAATGCCAAGGGGGGAAGGATGCCCTGTC	741
637	Db	GATTTTCCCATTTGCCATGATGGCATGCTGTGTGCTGGAAATACCGAGAGAGATCTCTGC	696
742	QY	TTCGGTGACTCAGGTGGACCTTGGCCCTGTAAACAAGAAATGACCTGTGGTATTCAGATTGGA	801
697	Db	CAGGGCGAATTCAGGGGGGCCACTGGCTGTGCAAAAGTGAAGGATACCTGGCTGCAGGCAGGA	756
802	QY	GTCGTGAGCTGGGAGTGGGCTGTGGTTCGGCCCCAATCGGCCCGGTGTCTACACCAATATC	861
757	Db	GTGCTCAGCTGGGTGAGGGCTGGCGACAGCCCAACAAAGCCTGGCATCTACACCCGGGTG	816
862	QY	AGCCACACATTTGAGTGGATCCA	884
817	b	ACATACTACTTAGACTGGATCCA	839

```

RESULT 9
; US-9-386-642-8
; Sequence 8, Application US/093866642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: QI, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/37-31
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of
; OTHER INFORMATION: with homo sapien
US-9-386-642-8

```

Query Match 10.9%; Score 117.8; DB 4; Length 1142;

Best Local Similarity 54.7%; Pred. No. 1.8e-21;  
Matches 261; Conservative 0; Mismatches 207; Indels 9; Gaps 1;

QY	429	CTGTGGGAATTCACCCCTATGACATTTGGCTTGGTGAAGCTGTCTCGACCTGTGCACCTACAC	488
DB	387	CCAGGAGGGCTCCCGAGGCGACATTTGACCTCTCACTCAGCAGACCCCATCACCCTTCCTC	446
QY	489	TAAACACATCAGCCCATCTGTCTCCAGGGCTCCACATTTGAGTTTGAGAACGGGACAGA	548
DB	447	CCGCTACATCCGGGCCCATCTGCCTCCCTGCAGCCAAAGCCCTCTCCCAAGCGCCTCCA	506
QY	549	CTGTGGGTGACTGGCTGGGGGTACATCAAAAGAGGATGAGGCATGCCATCTCCCCACAC	608
DB	507	CTGCACTGTCTACTGGCTGGGGTCATGTGGCCCCCTCAGTGAGCCTCCTGACGCCCAAGCC	566
QY	609	CCTCCAGGAACTTCAGTCCGCATCATAAACAACCTATGTGCACCACTCTTCTCTCAA	668
DB	567	ACTCCACAACCTCGAGGTGCTCTGATCATGTCTGAGAGCTGTAACTGCTGTACAACAT	626
QY	569	GTACAGTTTCCGCAAGGA-----CATCTTTGGAGACATGTTGTGCTGGCAATGC	719
DB	627	CGAGCCCAAGCCTGAGGAGCGGCACATTTGTCCAAGAGACATGGTGTGTCTGGGTATGT	686
QY	720	CCAAGGCGGGAAGGATTCCTGCTTCGGTCACTCAGGTGGACCCCTTGGCCTGTAAACAAG	779
DB	687	GGAGGGGGCAAGGACCCCTGCAGAGGTGACTCTGGGGGGCCCACTCTCTGCGCCTGTGA	746
QY	780	TGCACTGTGGTATCAGATTGAGTCTGAGCTGGGGAGTGGCCTGTGTGCGGCCCAATCG	839
DB	747	GGGTCTCTGGTACCTGACGGGCAATTTGTAGCTGGGGAGATGCCCTGTGGGGCCCCACAG	806
QY	840	GCCCGTGTCTACACCAATATCAGCCACCATTTTCAGTTGGATCAGAAAGCTGATGCC	896
DB	807	GCCGTGGTGTACACTCTGTGGCCTCCAGCTATGCGCTCTCTGGATCCAAAGCAAGGTGAC	863

```

RESULT 10
US-09-386-642-7
; Sequence 7, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: QI, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation
; FILE REFERENCE: ORN-1028
; CURRENT APPLICATION NUMBER: US/09/7
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of
; -OTHER INFORMATION: with homo sapien
US-09-386-642-7

```

Query Match 10.9%; Score 117.8; DB 4; Length 1169;  
Best Local Similarity 54.7%;  
Pred. NO. 1.9e-21;  
Matches 261; Conservative 0; Mismatches 207; Indels 9; Gaps 1;

Qy	429	CCTGGGAAATCACCCTATGACATTTGCCTTGGTGAAGCTCTCTGCACCTGTGCACCTACAC	488
Db	414	CCAGAGGGTCCCGAGGGGACATATGCACTCTCCAACTCAGCAGACCATCACCTCTTC	473
Qy	489	TAAACATCATCGCCCATCTCTCCAGGCCTCCACATTTGATTTGAAACCGGACAGA	548
Db	474	CCGCTACATCCGGGCCATCTGCGCTCCCTCGACGCAACGCGCTCTTCCCCAAGGGCTCCA	533
Qy	549	CTGCTGGGTGACTTGCTGGGGGTACATCAAGAGGATGAGGCACCTGCCCATCTCCCCACAC	608

Db 534 CTGCACTGTCTACTGGTGGGTCTATGTGGCCCCCTCACTGAGCCTCTCTGAGGCCCAAGCC 593  
 Qy 609 CCTCAGGAAGTCTCAGGTGCGCATATAACAACACTCTATGTGCAACCACTCTCTCTCAA 668  
 Db 594 ACTGAGCAACTCAGGTGCGCTCTGATCAGTCGTGAGACGTTGTAATCTCTCTGTAACAT 653  
 Qy 669 GTACAGTTTCGCAAGGA-----CATCTTTGGAGACATGTTTGTGCTGCGCAATGC 719  
 Db 654 CGAGCGCAAGCTCAGGAGCGCACTTTGTCGAAGAGGACATGTTGTGCTGGCTATGT 713  
 Qy 720 CCAGGGGGGAAGTACCTCTCTCGTGAATCAGGTGAGCCCTGGCTGTAAACAAGAA 779  
 Db 714 GGAGGGGGGAAGAGCGCTCTCAGGTGAGTCTGGGGGCCCACTCTCTCTGCTGTTGA 773  
 Qy 780 TGGAGTGTGTTATCAGATTGAGTCTGAGTGGGGAGTGGCTGTGTCGCGCCCAATCG 839  
 Db 774 GGGTCTGTGTACCTGAGCGCAATGTGAGTGGGGAGATCCCTGTGGGGCCCGACAG 833  
 Qy 840 GCGCGGTGTACACCAATATACGCCCACTTTGAGTGGATCCAGAGCTGATGGC 896  
 Db 834 GCGTGGTGTACACTCTGGCCTCCAGCTATGCTCTGATCCAAAGAGGTGAC 890

RESULT 11

US-09-016-366A-18  
 ; Sequence 18, Application US/09016366A  
 ; Patent No. 5955431  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stevens, Richard L.  
 ; APPLICANT: Huang, Chifu  
 ; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
 ; TITLE OF INVENTION: INHIBITORS  
 ; NUMBER OF SEQUENCES: 65  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 ; STREET: 600 Atlantic Avenue  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02210-2211  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette.  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/016,366A  
 ; FILING DATE: January 30, 1998  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/037,090  
 ; FILING DATE: 05-FEB-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Plumer, Elizabeth R.  
 ; REGISTRATION NUMBER: 36,637  
 ; REFERENCE/DOCKET NUMBER: B0801/7093  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-720-3500  
 ; TELEFAX: 617-720-2441  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1137 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cdna  
 ; US-09-016-366A-18

Query Match 10.88; Score 117; DB 2; Length 1137;  
 Best Local Similarity 55.88; Pred. No. 3e-21;  
 Matches 247; Conservative 0; Mismatches 190; Indels 6; Gaps 1;

Qy 448 GACATTTGCTTGGTGAAGCTGTCTGCACCTGTCACTTACCTAAACACATCCAGCCCATC 507  
 Db 357 GACATCGCCCTGCTGGAGCTGAGGAGCGGTGACGCTCCAGGCCACGTCACACGGTC 416  
 Qy 508 TGTCTCCAGGCTTCACATTTGAGTTTGAACCCGACAGACTGCTGGTGACTGGCTGG 567  
 Db 417 ACCCTGCCCCCTGCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 476  
 Qy 568 GGGTACATCAAGAGGATGAGGCACTGCATCTCCCAACACCTCCAGGAAGTTTCAGTGC 627  
 Db 477 GCGATGTGGACATGATGAGGCTCCACCGCCATTTCTCTGAACAGGTTGAGGTC 536  
 Qy 628 GCCATATTAACAACCTCTATGTG-----CAACCACTCTTCTCTCAAGTACAGTTTCGCG 681  
 Db 537 CCCATAATGGAACACCATTTGTGACGCAAAATACCACCTTGGCGCTCACACGGAGAC 596  
 Qy 682 AAGNACATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGGGGGAAGATCCCTGCG 741  
 Db 597 GACGTCCCATCGTCCGTCGACACATGCTGTGTCGGGGAACACCCGAGGGGACTCATGC 656  
 Qy 742 TTCGTGACTCAGGTGGACCTTGGCCCTTAACAGAATGGACTGTGTATCAGATTGGA 801  
 Db 657 CAGGCGACTCCGAGGCGCCCTGCTGTGCAAGGTGAATGGACCTGCTGCGAGCGGCG 716  
 Qy 802 GTCGTGAGCTGGGAGTGGGTGTGTCGCGCCCAATCGCCCGGTGTCTACACCAATATC 861  
 Db 717 GTGTGACGTGGGCGAGGCTGTGCCAGCCCAACCGGCTGTCATCTACACCGGTGTC 776  
 Qy 862 AGCCACCACTTTGAGTGGATCCA 884  
 Db 777 ACCTACTACTTGGACTGGATCCA 799

RESULT 12

US-08-978-404B-13  
 ; Sequence 13, Application US/08978404B  
 ; Patent No. 5968782  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stevens, Richard L.  
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
 ; TITLE OF INVENTION: FIBRINOGEN  
 ; NUMBER OF SEQUENCES: 74  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
 ; STREET: 600 Atlantic Avenue  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02210-2211  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/978,404B  
 ; FILING DATE: 25-NOV-97  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/032,354  
 ; FILING DATE: 04-DEC-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Plumer, Elizabeth R.  
 ; REGISTRATION NUMBER: 36,637  
 ; REFERENCE/DOCKET NUMBER: B0801/7090  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-720-3500  
 ; TELEFAX: 617-720-2441  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 13:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1137 base pairs  
 ; TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-978-404B-13

Query Match 10.8%; Score 117; DB 2; Length 1137;  
Best Local Similarity 55.8%; Pred. No. 3e-21;  
Matches 247; Conservative 0; Mismatches 190; Indels 6; Gaps 1;

QY 448 GACATTGCCTTGTGAAGCTGTCTGACCTGTCTACCTACACTAAACACATCCAGCCCATC 507  
DB 357 GACATCGCCCTGTGTGAGCTGGAGGAGCGGTGAAGCTCTCCAGCCACGTCACACGGTC 416  
QY 508 TGTCTCCAGGCCCTCCACATTTGAGTTTGAGAACCGGACACAGACTGTGGTGTGACTGG 567  
DB 417 ACCTCGCCCTCCCTCAGAGACCTTCCCCCGGGGATCCGTGTGGTGTCACTGGCTGG 476  
QY 568 GGTATCATCAAGAGAGGAGGACCTGCCATCTCCACACACCTCCAGGAATTCAGTTC 627  
DB 477 GCGGATGTGACAAATGATGAGCGCCTCCACCGCCATTTCCCTCTGAAGCAGGTGAAGTTC 536  
QY 628 GCATCATFAAACAACTCTATGTG-----CAACCACTCTTCTCTCAAGTACAGATTCCGC 681  
DB 537 CCCCATAATGGAACACATTTGTGACGCAAAATACCACTTGGCGCTACACGGGAGAC 596  
QY 682 AAGGACATCTTTGGAGACATGTTTGTCTGCTGCAATGCCAAGGCGGGAAGGATCCCTGC 741  
DB 597 GAGTCCCGATCTCCGTGACGACATGCTGTGTGCCCGGGAACACCCGAGGAGCTCATGC 656  
QY 742 TTCGGTACTCAGTGGAGCCCTTGGCTGTGTAACAAAGATGGACTGTGTATCAGATTGGA 801  
DB 657 CAGGCGACTCGGAGGCGCCCTGTGTGCAAGGTGAATGGACTGTGCTGCGCGGGC 716  
QY 802 GTCTGTAGCTGGGAGTGGCTGTGTGCGGCCCAATCGGCCCGGTGTCTACACCAATATC 861  
DB 717 GTGGTGTAGCTGGGAGGCGGTGTGCCAGCCCAACCGGCTGCGCATCTACACCGGTGC 776  
QY 862 AGCCACCACTTTGAGTGGATCCA 884  
DB 777 ACCTACTACTTGGACTGGATCCA 799

## RESULT 13

US-09-510-738A-188  
; Sequence 188, Application US/09510738A  
; Patent No. 6268165  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of  
; FILE REFERENCE: D6223CIP-A  
; CURRENT APPLICATION NUMBER: US/09/510,738A  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 09/039,211  
; NUMBER OF SEQ ID NOS: 188  
; SEQ ID NO 188  
; LENGTH: 1783  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: full length cDNA of hepsin  
US-09-510-738A-188

Query Match 10.7%; Score 116.2; DB 4; Length 1783;  
Best Local Similarity 51.0%; Pred. No. 5.5e-21;  
Matches 398; Conservative 0; Mismatches 353; Indels 30; Gaps 4;

QY 114 GCGCGCGCCCTTATCAGGACCATCGCGCGGCGGTCTACGTGCGCATCGTGGTGG 173  
DB 683 GCGCGCCATCTGCCAAGACTGTGGCGCAGGAAGTCCCGTGGACCGCATCGTGGAGG 742  
QY 174 AGAGACGCCGAACATCGGCGCTGTGGCGTGGCAGGGAGCTGGCCCTGTGGGATCCCA 233

Query Match 10.7%; Score 116.2; DB 4; Length 2363;

DB 743 CCGGACACACAGCTTGGCGCGGTGGCGAAGTCAGCCTTCGCTATGATGAGACACA 802  
QY 234 CGTATGGGAGTGAGCCTGCTCAGCCACCGCTGGGCACTCACGCGCGGCACTGCTTTGA 293  
DB 803 CTTCTGTGGGGATCCTGCTCTCCGGGACTGGGTGCTCAGACCGCCCACTGCTT--- 859  
QY 294 AACCTATAGTACCTTAGTGTATCCCTCCGGGTGGATGCTCCAGCTTTGGCCAGCTGACTTC 353  
DB 860 CCGGAGCGAAGACCGGCTCCTGTCCCGATGGGAGTGTTCGCGGTGCGGTGCCCGCAGC 919  
QY 354 CATCGCATCTCTTGTGAGCCTGACGCTTACTACACCCGTTACTTCTGATATGAAATATCTA 413  
DB 920 CTTCTCCCAAGCTCTGACCTGGGGTGCAGGTGTGGTCTACCAAGGGGGTATCTTCC 979  
QY 414 TCTGAGCCCTCCCTACCTACCTGGGAATTCACCTATGACATTTGCTTGGTGAAGTGTCTGC 473  
DB 980 CTTTGGGAGCCCAACAGCGAGGAGAACAGCAAGATATTCCTGCTGCTCCACTCTCCAG 1039  
QY 474 ACCTGTACCTACACTAACACATCCAGCCATCTGTCTCCAGGCTCCACATTTGAGTT 533  
DB 1040 TCCCTGCCCTTCACAGAAATACATCCAGCTGTGTCTCCAGCTGCCGGCCAGGCCCT 1099  
QY 534 TGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGTACATCAAGAGGATGAGCACT 593  
DB 1100 GTGGATGCAAGATCTGTACCGTACCGGCTGGGGCAACACGCACTACTATGGCCAACA 1159  
QY 594 GCCATCTCCCAACCCCTCCAGGAAGTTTCAGGTGCGCCATCAAAACACTCTATGTGCA 653  
DB 1160 G-----GCCGGGTACTTCCAGAGGCTCGAGTCCCAATATCAGCAATGATGTCTGCA 1213  
QY 654 CCACCTCTTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGTTTGTGTGG 713  
DB 1214 TGGCGCTGACTTC-----TATGAAACCATGATCAAGCCCAAGATGTTCTGTGTGG 1264  
QY 714 CAATGCCCAAGCGGGAGGATGCTTCTCGGTGACTCAGGTGAGCCCTTGGCCCTGTAA 773  
DB 1265 CTACCCGAGGTGCAATGATGCTCCAGGCGCAGCGGTGTGCTTGTGTGTGA 1324  
QY 774 CAA-----GAATGACTGTGTATCAGATTGGAGTCTGAGCTGGGAGTGGG 821  
DB 1325 GGAGACATCTCTCGGACGCGCAGCTTGGCGCTGTGTGGCATTTGAGTTGGGCACTGG 1384  
QY 822 CTGTGTGCGCCCAATCGGCGCGGTGCTACACCAATATCAGCCACCATTTGAGTGGAT 881  
DB 1385 CTGTGCTGCGCCAGCAAGCCAGCGGTCTACACCAAGTCAGTCTCCGGGAGTGGAT 1444  
QY 882 C 882  
DB 1445 C 1445

## RESULT 14

US-09-742-703-3  
; Sequence 3, Application US/09742703  
; Patent No. 6423543  
; GENERAL INFORMATION:  
; APPLICANT: Patrick Allen Marcotte  
; APPLICANT: Lex M. Cowsett  
; TITLE OF INVENTION: ANTISENSE MODULATION OF HEP SIN EXPRESSION  
; FILE REFERENCE: RTS-0090  
; CURRENT APPLICATION NUMBER: US/09/742,703  
; CURRENT FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 49  
; SEQ ID NO 3  
; LENGTH: 2363  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (826)..(2079)  
US-09-742-703-3

Query Match

Best Local Similarity 51.0%; Pred. No. 6e-21;  
Matches 398; Conservative 0; Mismatches 353; Indels 30; Gaps 4;

QY 114 GCGCGGCGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 173  
Db 1263 GCGCGGCGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1322

QY 174 AGAGGACGCGGCGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 233  
Db 1323 CCGGACACAGCTTGGGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1382

QY 234 CATTGCGGAGTGGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 293  
Db 1383 CCGGAGGAGTGGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1439

QY 294 AACCTATAGTGGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 353  
Db 1440 CCGGAGGAGTGGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1499

QY 354 CATGCCATCTTCTGAGGCGCTGAGGCGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 413  
Db 1500 CTCTCCGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1559

QY 414 TCTGAGCGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 473  
Db 1560 CTTTCGGGACCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1619

QY 474 ACCTGTCACCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 533  
Db 1620 TCCCGTCCGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1679

QY 534 TGAGAACCGGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 593  
Db 1680 GGTGATGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1739

QY 594 GCGATCTCCGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 653  
Db 1740 G-----GCCGGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1793

QY 654 CCACCTCTCTCAAGTACGCTTCCGCAAGGACATCTTGGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 713  
Db 1794 TGGCGCTGCTTCTCAAGTACGCTTCCGCAAGGACATCTTGGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1844

QY 714 CAATGCCGCGGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 773  
Db 1845 CTACCCGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1904

QY 774 CAA-----GAATGGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 821  
Db 1905 GCGACGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 1964

QY 822 CTGCGTCCGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 881  
Db 1965 CTGCGTCCGCGGCGGCTTATCAGGACCATGCGCGGCGGCGGCTATCAGCTGCGGCATGCTGGGCG 2024

QY 882 C 882  
Db 2025 C 2025

RESULT 15  
US-09-079-970A-1  
; Sequence 1, Application US/09079970A  
; Patent No. 6274366  
; GENERAL INFORMATION:  
; APPLICANT: Maffitt, Mark A.  
; APPLICANT: Niles, Andrew L.  
; APPLICANT: Haak-Frendscho, Mary  
; TITLE OF INVENTION: Enzymatically-Active Recombinant Human  
; TITLE OF INVENTION: Beta-Tryptase and Method of Making Same  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Intellectual Property Department

STREET: 8000 Excelsior Drive, Suite 401  
CITY: Madison  
STATE: WISCONSIN  
COUNTRY: U.S.A.  
ZIP: 53717-1914  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,970A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Leone, Joseph T.  
REGISTRATION NUMBER: 37,170  
REFERENCE/DOCKET NUMBER: 34506.073  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 831-2100  
TELEFAX: (608) 831-2106  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..735  
US-09-079-970A-1

Query Match 10.7%; Score 115.4; DB 4; Length 735;  
Best Local Similarity 55.5%; Pred. No. 6.7e-21;  
Matches 246; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

QY 448 GACATTCGCTTGGTGAAGCTGTCTGCACCTGTGCACCTACACTAAACACATCCAGGCCATC 507  
Db 271 GACATTCGCTTGGTGAAGCTGTCTGCACCTGTGCACCTACACTAAACACATCCAGGCCATC 330

QY 508 TGTCTCCAGGCTCCACATTTGAGTTGAGACCGGACAGACTGCTGGGTGACTGCTGCTG 567  
Db 331 ACCCTGCGGCTTGGTGAAGCTGTCTGCACCTGTGCACCTACACTAAACACATCCAGGCCATC 390

QY 568 GGTATCATCAAGAGATGAGGACCTGCTCCACCTCCACACACCTCCAGGAAGTTGAGTTC 627  
Db 391 GCGGATGAGCAATGATGAGGCGCTCCACCGCCATTTCTCTGAAGCAGGTGAAGTTC 450

QY 628 GCCATCAATAACACTCTATGTC-----CAACACACTCTTCTCTCAAGTACAGTTTCCGC 681  
Db 451 CCCATAATGGAACACATTTGTGACGCAAAATACCACTTTGGGCGCTTACACGGGAGAC 510

QY 682 AAGGACATCTTGGAGACATGTTGTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 741  
Db 511 GACGTCCGCTTGGAGACATGTTGTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 570

QY 742 TTCGGTACTCAGGTGGACCTTGGCTTAAACAAGATGGACTGTGGTATCAGATTGGA 801  
Db 571 CAGGCGGACTCCGAGGCGGCTTGGTGTGCAAGGTGATGGACCTTGGTGTGAGGCGGCG 630

QY 802 GTCGTGAGTGGGAGTGGGCTGTGTGTCGCGCCCAATTCGCGCGGCTGTCTTACACCAATATC 861  
Db 631 GTGGTCAAGTGGGAGGCTGTGTCGCAAGGTGATGGACCTTGGTGTGAGGCGGCG 690

QY 862 AGCCACCACTTTGAGTGGATCCA 884  
Db 691 ACCTACTTGGACTGGATCCA 713

Mon May 12 16:37:39 2003

Search completed: May 10, 2003, 04:22:18  
Job time : 80 secs

us-09-787-844-1.rni

Page 10

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 02:58:22 ; Search time 3359 Seconds  
(without alignments)  
9365.922 Million cell updates/sec

Title: US-09-787-844-1  
Perfect score: 1081  
Sequence: 1 agaggcagagggggcgtca.....acattccagttgatgcctgc 1081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*\*  
1: gb\_ba:\*\*  
2: gb\_htg:\*\*  
3: gb\_in:\*\*  
4: gb\_om:\*\*  
5: gb\_ov:\*\*  
6: gb\_pat:\*\*  
7: gb\_ph:\*\*  
8: gb\_pl:\*\*  
9: gb\_pr:\*\*  
10: gb\_ro:\*\*  
11: gb\_sts:\*\*  
12: gb\_sy:\*\*  
13: gb\_un:\*\*  
14: gb\_vi:\*\*  
15: em\_ba:\*\*  
16: em\_fun:\*\*  
17: em\_hum:\*\*  
18: em\_in:\*\*  
19: em\_mu:\*\*  
20: em\_om:\*\*  
21: em\_or:\*\*  
22: em\_ov:\*\*  
23: em\_pat:\*\*  
24: em\_ph:\*\*  
25: em\_pl:\*\*  
26: em\_ro:\*\*  
27: em\_sts:\*\*  
28: em\_un:\*\*  
29: em\_vi:\*\*  
30: em\_htg\_hum:\*\*  
31: em\_htg\_inv:\*\*  
32: em\_htg\_other:\*\*  
33: em\_htg\_mus:\*\*  
34: em\_htg\_pln:\*\*  
35: em\_htg\_rod:\*\*  
36: em\_htg\_mam:\*\*  
37: em\_htg\_vrt:\*\*  
38: em\_sy:\*\*  
39: em\_htgo\_hum:\*\*  
40: em\_htgo\_mus:\*\*  
41: em\_htgo\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1081	100.0	1081	6	AR142617	AR142617 Sequence
2	1054.8	97.6	1077	6	AX370460	AX370460 Sequence
3	1054.8	97.6	1077	9	AF058300	AF058300 Homo sapi
4	1045.8	96.7	1082	6	AX370459	AX370459 Sequence
5	1045.8	96.7	1082	6	E21866	E21866 Novel acido
6	1045.8	96.7	1082	6	E21867	E21867 Novel acido
7	1045.8	96.7	1082	9	AB031329	AB031329 Homo sapi
8	1039.8	96.2	1071	9	AB031330	AB031330 Homo sapi
9	1034.4	95.7	1085	6	E21853	E21853 Novel acido
10	1034.4	95.7	1085	6	E21854	E21854 Novel acido
11	969.2	89.7	1119	9	AB031331	AB031331 Homo sapi
12	938.8	86.8	942	6	E21865	E21865 Novel acido
13	927.4	85.8	945	6	E21852	E21852 Novel acido
14	493.4	45.6	529	6	AX321357	AX321357 Sequence
15	489	45.2	1055	10	AY005145	AY005145 Mus muscu
16	489	45.2	1084	10	AB033822	AB033822 Mus muscu
17	489	45.2	1095	10	AF176209	AF176209 Mus muscu
18	487.8	45.1	1063	10	AB059415	AB059415 Mus muscu
C 19	465.4	43.1	575	6	AX198283	AX198283 Sequence
C 20	465.4	43.1	575	6	AX208847	AX208847 Sequence
21	422.2	39.1	888	6	AX360070	AX360070 Sequence
22	413.4	38.2	1018	6	AX370454	AX370454 Sequence
23	338.6	31.3	1128	10	AB049453	AB049453 Mus muscu
24	334.2	30.9	5153	9	AF058301	AF058301 Homo sapi
C 25	334.2	30.9	35677	9	AC005361	AC005361 Homo sapi
C 26	332.6	30.8	199706	2	AC097329	AC097329 Pan trogl
C 27	329.8	30.5	837	6	AX370461	AX370461 Sequence
C 28	300	27.8	211962	9	AC092117	AC092117 Homo sapi
29	166.4	15.4	5015	10	AF226710	AF226710 Mus muscu
30	166.4	15.4	5551	10	AF304012	AF304012 Mus muscu
31	166.4	15.4	5628	10	AB059414	AB059414 Mus muscu
32	166.4	15.4	17405	10	AB041645	AB041645 Mus muscu
33	165	15.3	1013	9	HSA306593	HSA306593 Homo sapi
34	165	15.3	1077	9	AY030095	AY030095 Homo sapi
35	165	15.3	1129	6	AX464088	AX464088 Sequence
36	164.8	15.2	205939	2	AC110262	AC110262 Mus muscu
37	161.6	14.9	2457	6	AX360096	AX360096 Sequence
38	161.6	14.9	2681	6	AX480935	AX480935 Sequence
39	160	14.8	1488	10	BC024903	BC024903 Mus muscu
40	155.8	14.4	2066	9	AK055576	AK055576 Homo sapi
41	153	14.2	1259	10	RRAS642	AJ005642 Rattus ra
42	150.6	13.9	1199	9	AF175522	AF175522 Homo sapi
43	150.6	13.9	2662	6	AX338508	AX338508 Sequence
44	149	13.8	1958	6	AX342632	AX342632 Sequence
45	148.2	13.7	199706	2	AC097329	AC097329 Pan trogl

ALIGNMENTS

RESULT 1  
AR142617  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AR142617  
Sequence 15 from patent US 6203979.  
AR142617  
AR142617.1 GI:15103903  
Unknown.  
Unknown.  
Unclassified.  
1 (bases 1 to 1081)  
Bandman, O., Hillman, J.L., Yue, H., Guegler, K.J., Corley, N.C.,  
Tang, Y. Tom. and Shah, P.  
Human protease-molecules  
Patent: US 6203979-A 15 20-MAR-2001;

1081 bp  
DNA  
linear  
PAT 08-AUG-2001





QY 502 CCCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGACT 561  
Db 481 CCCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGACT 540  
QY 562 GGCTGGGGTACATCAAGAGGATGAGGACTGCCATCTCCCAACACCTCCAGGAAGTT 621  
Db 541 GGCTGGGGTACATCAAGAGGATGAGGACTGCCATCTCCCAACACCTCCAGGAAGTT 600  
QY 622 CAGTCCGCCATCAATAACAACCTTATGTGAACACACCTCTTCCCTCAAGTACAGTTCCGC 681  
Db 601 CAGTCCGCCATCAATAACAACCTTATGTGAACACACCTCTTCCCTCAAGTACAGTTCCGC 660  
QY 682 AAGACATCTTTGAGACATGTTTGTGGGCAATGCCAAGCGGGAAGATGCTGTC 741  
Db 661 AAGACATCTTTGAGACATGTTTGTGGGCAATGCCAAGCGGGAAGATGCTGTC 720  
QY 742 TTCGGTACACTGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 801  
Db 721 TTCGGTACACTGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 780  
QY 802 GTCGTGAGCTGGGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 861  
Db 781 GTCGTGAGCTGGGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 840  
QY 862 AGCCACCACCTTTGAGTGGATCCAGAACTGATGCCCGGAGTGGCATGTCGCCAGAC 921  
Db 841 AGCCACCACCTTTGAGTGGATCCAGAACTGATGCCCGGAGTGGCATGTCGCCAGAC 900  
QY 922 CCTCTCGGACACTACTCTTTTCCCTTCTCTGCGGCTCTCCACACTCTCGGGCCGGTC 981  
Db 901 CCTCTCGGACACTACTCTTTTCCCTTCTCTGCGGCTCTCCACACTCTCGGGCCGGTC 960  
QY 982 TGAGCTTACCTGAGCCACTGAGGCTGGGCGCACTGCAAGTCAAGGCTGCTCTCTTC 1041  
Db 961 TGAGCTTACCTGAGCCACTGAGGCTGGGCGCACTGCAAGTCAAGGCTGCTCTCTTC 1020  
QY 1042 TGCTCTTTGGTAATAACACATTCAGTTGATGCT 1079  
Db 1021 TGCTCTTTGGTAATAACACATTCAGTTGATGCT 1058

RESULT 3  
AF058300  
LOCUS Homo sapiens testisin (TEST1) mRNA, complete cds. PRI 01-JUL-1999  
ACCESSION AF058300  
VERSION AF058300.1 GI:5305322  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Hooper, J.D., Dickinson, J.L. and Antalis, T.M.  
TITLE Direct Submission  
JOURNAL Submitted (08-APR-1998) Queensland Institute of Medical Research,  
300 Herston Road, Brisbane, QLD 4006, Australia  
FEATURES  
source  
1..1077  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16p13.3"  
/cell\_line="HeLa S2"  
1..1077  
/gene="TEST1"  
19..963  
/gene="TEST1"  
/note="serine proteinase"  
/codon\_start=1  
/product="testisin"  
/protein\_id="AAD41588.1"  
/db\_xref="GI:5305323"

/translation="MGARGALLALLALLARAGLRKRPESQEAAPLSPGCGRRVTSRIVG  
GDEALGRWPQGLSLRWDSHVGVSLLSHRWALTAHCFETYSDLSDPGWMVQFQ  
LTMSPFWSQAVYTRFVSNIYLSPIYGNVPIALVILSAPVTYTKIOPICLOA  
STFEENRTDCWTGMYIKEDRALPSPHLOEVOVAI INNSMCHLEFLKVSFRKDF  
GDVACNAAGGKDACDAGSGGLACNKNGLWYIGVSWGVGGRNRPRVPTNISH  
HFWDIOKLMAQSGMSQDPSPWPLLLFFLLWALPLLGPV"  
polya\_signal 1034..1039  
/gene="TEST1"  
BASE COUNT 202 a 320 c 314 g 241 t  
ORIGIN  
Query Match 97.6%; Score 1054.8; DB 9; Length 1077;  
Best Local Similarity 99.8%; Pred. No. 7e-236;  
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 22 GCCCGGGAGAGGAGCCATGGCGCGCGGGGCTGCTGCGCGCTGCTGCTGCT 81  
Db 1 GCCCGGGAGAGGAGCCATGGCGCGCGGGGCTGCTGCTGCTGCTGCTGCT 60  
QY 82 CGGCTGAGCTCAGGAAGCCGAGTTCGAGGAGGCGGCGCTTATCAGGACCATGCGGC 141  
Db 61 CGGCTGAGCTCAGGAAGCCGAGTTCGAGGAGGCGGCGCTTATCAGGACCATGCGGC 120  
QY 142 CGACGGGTATCATCGTCCGCGCATCGTGGTGGAGGAGACGCCAATCGGGGCTGGCG 201  
Db 121 CGACGGGTATCATCGTCCGCGCATCGTGGTGGAGGAGACGCCAATCGGGGCTGGCG 180  
QY 202 TGGCAGGGAGCCCTGGCGCTGCGGATCCACAGTATGCGGAGTCAAGCTGCTCAGCCAC 261  
Db 181 TGGCAGGGAGCCCTGGCGCTGCGGATTCACAGTATGCGGAGTCAAGCTGCTCAGCCAC 240  
QY 262 CGCTGGGCACTCAGCGGGCGCACTGCTTTGAAACCTATAGTACCTTAGTATGATCCCTCC 321  
Db 241 CGCTGGGCACTCAGCGGGCGCACTGCTTTGAAACCTATAGTACCTTAGTATGATCCCTCC 300  
QY 322 GGGTGGATGTCAGTTTGGCCAGTTCATGCTCCATGCCATCCTCTGAGGCTTGAGGCC 381  
Db 301 GGGTGGATGTCAGTTTGGCCAGTTCATGCTCCATGCCATCCTCTGAGGCTTGAGGCC 360  
QY 382 TACTACACCCGCTTACTTCTGTAATATCTATCTGAGCCCTCGCTACTCGGGAATCA 441  
Db 361 TACTACACCCGCTTACTTCTGTAATATCTATCTGAGCCCTCGCTACTCGGGAATCA 420  
QY 442 CCCTATGACATTTGCTGGTGAAGCTGCTGCACCTGTACCTACATTAACACATCAG 501  
Db 421 CCCTATGACATTTGCTGGTGAAGCTGCTGCACCTGTACCTACATTAACACATCAG 480  
QY 502 CCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGACT 561  
Db 481 CCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGACT 540.  
QY 562 GGCTGGGGTACATCAAGAGGATGAGGACTGCCATCTCCCAACACCTCCAGGAAGTT 621  
Db 541 GGCTGGGGTACATCAAGAGGATGAGGACTGCCATCTCCCAACACCTCCAGGAAGTT 600  
QY 622 CAGTCCGCCATCAATAACAACCTTATGTGAACACACCTCTTCCCTCAAGTACAGTTCCGC 681  
Db 601 CAGTCCGCCATCAATAACAACCTTATGTGAACACACCTCTTCCCTCAAGTACAGTTCCGC 660  
QY 682 AAGACATCTTTGAGACATGTTTGTGGGCAATGCCAAGCGGGAAGATGCTGTC 741  
Db 661 AAGACATCTTTGAGACATGTTTGTGGGCAATGCCAAGCGGGAAGATGCTGTC 720  
QY 742 TTCGGTACACTGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 801  
Db 721 TTCGGTACACTGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 780  
QY 802 GTCGTGAGCTGGGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 861  
Db 781 GTCGTGAGCTGGGAGTGGGCTGTGTGGCCCAATCGGCCGCTGTCTACACCAATATC 840  
QY 862 AGCCACCACCTTTGAGTGGATCCAGAACTGATGCCCGGAGTGGCATGTCGCCAGAC 921  
Db 841 AGCCACCACCTTTGAGTGGATCCAGAACTGATGCCCGGAGTGGCATGTCGCCAGAC 900

Db 841 AGCCACCATTGAGTGGATCCAGAAGCTGATGGCCAGAGTGCGATGTCCAGCCAGAC 900

QY 922 CCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGGCGGTC 981

Db 901 CCCTCTGGCGGCTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGGCGGTC 960

QY 982 TGAGCCTACTGAGCCATGCGAGCCTGGGGCCACTGCCAAGTCAGGCCCTGTTCTCTTC 1041

Db 961 TGAGCCTACTGAGCCATGCGAGCCTGGGGCCACTGCCAAGTCAGGCCCTGTTCTCTTC 1020

QY 1042 TGTCTTGTGTTGTAATAAACACATCCACGTTGATGCT 1079

Db 1021 TGTCTTGTGTTGTAATAAACACATCCACGTTGATGCT 1058

RESULT 4

AX370459

LOCUS

AX370459 1082 bp DNA linear PAT 16-FEB-2002

DEFINITION

Sequence 6 from Patent WO0198503.

ACCESSION

AX370459

VERSION

AX370459.1 GI:18857513

KEYWORDS

human.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Xiao,Y.

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE

Regulation of human eosinophil serine protease 1-like enzyme

JOURNAL

Patent: WO 0198503-A 6 27-DEC-2001;

FEATURES

Location/Qualifiers

1..1082

source

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT

216 a 318 c 308 g 240 t

ORIGIN

Query Match 96.7%; Score 1045.8; DB 6; Length 1082;

Best Local Similarity 99.8%; Pred. No. 8.9e-234;

Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 31 GAGGAGCCATGGCGCGCGCGGCGCTGCTGCTGGCGCTGCTGGCTGGCGCTGGA 90

Db 1 GAGGAGCCATGGCGCGCGCGGCGCTGCTGCTGGCGCTGCTGGCTGGCGCTGGA 60

QY 91 CTCAGGAAGCCGAGTGCAGGAGGCGGCGCTTATCAGGACCATGCGGCGGCGGTC 150

Db 61 CTCAGGAAGCCGAGTGCAGGAGGCGGCGCTTATCAGGACCATGCGGCGGCGGTC 120

QY 151 ATCAGCTCGGCATGCTGGGTGGAGAGGAGCGGCAACTCGGCGTTGGCGTGGCAGGG 210

Db 121 ATCAGCTCGGCATGCTGGGTGGAGAGGAGCGGCAACTCGGCGTTGGCGTGGCAGGG 180

QY 211 AGCTCGGCTGTGGATCCACGATGCGGAGTGGAGTGGCTGCTGAGCCCTGAGCCGTC 270

Db 181 AGCTCGGCTGTGGATCCACGATGCGGAGTGGAGTGGCTGCTGAGCCCTGAGCCGTC 240

QY 271 CTCAGCGGCGGCTGCTTTGAACCTATAGTACCTTAGTACCTTGGCTGCGGCTGGATG 330

Db 241 CTCAGCGGCGGCTGCTTTGAACCTATAGTACCTTAGTACCTTGGCTGCGGCTGGATG 300

QY 331 GTCCAGTTGGCCAGCTGACTTCCATGCGATCCTTCTTGAGCCTGAGCCCTTACTACCC 390

Db 301 GTCCAGTTGGCCAGCTGACTTCCATGCGATCCTTCTTGAGCCTGAGCCCTTACTACCC 360

QY 391 CGTTACTTGGTATCAATATCTATCTGAGCCCTGCTACCTTGGGGAATTCACCTATGAC 450

Db 361 CGTTACTTGGTATCAATATCTATCTGAGCCCTGCTACCTTGGGGAATTCACCTATGAC 420

QY 451 ATGCGCTTGGTGAAGCTGTCTGACCTGTACCTACCTAACACATCCAGCCCATCTGT 510

Db 421 ATGCGCTTGGTGAAGCTGTCTGACCTGTACCTACCTAACACATCCAGCCCATCTGT 480

QY 511 CTCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGGTGGTGGTGGG 570

Db 481 CTCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGTGGTGGTGGTGGG 540

QY 571 TACATCAAGAGAGATGAGGACTGCCATCTCCCCACACCTCCAGGAAGTTTCAGGTCGCC 630

Db 541 TACATCAAGAGAGATGAGGACTGCCATCTCCCCACACCTCCAGGAAGTTTCAGGTCGCC 600

QY 631 ATCAATAACAACCTCTATGTGCAACACACCTCTTCCCTCAAGTACAGTTTCGCGAGGACATC 690

Db 601 ATCAATAACAACCTCTATGTGCAACACACCTCTTCCCTCAAGTACAGTTTCGCGAGGACATC 660

QY 691 TTTGAGAGATGTTTGTGCTGGCAATGCCCCAAGCGGGAAGTGCCTGCTTCGCTGAC 750

Db 661 TTTGAGAGATGTTTGTGCTGGCAATGCCCCAAGCGGGAAGTGCCTGCTTCGCTGAC 720

QY 751 TCAGGTGACCTTGGCTGTGTAAGAAATGAGTGTGTATCAGATTTGAGTTCGTCGAGC 810

Db 721 TCAGGTGACCTTGGCTGTGTAAGAAATGAGTGTGTATCAGATTTGAGTTCGTCGAGC 780

QY 811 TGGGAGTGGCTGTGCTGGCCCAATCGGCCCGCTGTCTACCAATATCAGCCACAC 870

Db 781 TGGGAGTGGCTGTGCTGGCCCAATCGGCCCGCTGTCTACCAATATCAGCCACAC 840

QY 871 TTTGAGTGGATCCAGAAAGTGTGATGCGCCAGAGTGGCATGTCCAGCCAGACCCCTCTGG 930

Db 841 TTTGAGTGGATCCAGAAAGTGTGATGCGCCAGAGTGGCATGTCCAGCCAGACCCCTCTGG 900

QY 931 CCCTACTCTTTTCCCTCTCTCTGGGCTTCCACTTCCGCGGCGGCTGAGCCTAC 990

Db 901 CCCTACTCTTTTCCCTCTCTCTGGGCTTCCACTTCCGCGGCGGCTGAGCCTAC 960

QY 991 CTGAGCCCATGAGCTTGGGCGCACTGCCAAGTCCAGCCCTGCTCTCTCTGCTTGT 1050

Db 961 CTGAGCCCATGAGCTTGGGCGCACTGCCAAGTCCAGCCCTGCTCTCTCTGCTTGT 1020

QY 1051 TGTAATAAACACATTTCCAGTTGATGCT 1079

Db 1021 TGTAATAAACACATTTCCAGTTGATGCT 1049

RESULT 5

E21866

LOCUS

E21866 1082 bp DNA linear PAT 18-JUN-2001

DEFINITION

Novel acidophil serine protease.

ACCESSION

E21866

VERSION

E21866.1 GI:13023737

KEYWORDS

JP 199032768-A/2.

SOURCE

unidentified.

ORGANISM

unclassified.

REFERENCE

1 (bases 1 to 1082)

AUTHORS

Hiroshi,K. and Masahiro,I.

TITLE

Novel acidophil serine protease

JOURNAL

Patent: JP 199032768-A 2 09-FEB-1999;

COMMENT

ONO PHARMACEUT CO LTD

OS Unidentified

PN JP 199032768-A/2

PD 09-FEB-1999

PF 16-JUL-1997 JP 1997191319

PR

PI HIROSHI KIDO,MASAHIRO INOUE

PC C12N15/09,A61K38/55,A61K39/395,A61K48/00,C07K7/00,

PC C07K16/40,

PC C12N9/64,C12N15/00,A61K37/64,A61K37/64

CC Strandedness: Single;

CC Topology: Linear;

FT key

FT source

FT Location/Qualifiers

1..1082

1..1082

Location/Qualifiers

/organism='Unidentified'.

source

1..1082

Qy	991	CTGAGCCCATGCAGCCTGGCGGCACATGCCAAGTCAGGCCCTGTGTCTCTTCTGTCTTTGTT	108
Db	961	CTGAGCCCATGCAGCCTGGCGGCACATGCCAAGTCAGGCCCTGTGTCTCTTCTGTCTTTGTT	102
Qy	1051	TGGTAATAAACACATTCCAGTTTGATGCCT	1079
Db	1021	TGGTAATAAACACATTCCAGTTTGATGCCT	1049
RESULT 6			
E21867	E21867	Novel acidophil serine protease.	
DEFINITION	E21867	Novel acidophil serine protease.	
ACCESSION	E21867	GI:13023738	
VERSION	E21867.1	JP 199032768-A/3.	
KEYWORDS	Homo sapiens.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1082)		
JOURNAL	Hiroshi,K. and Masahiro,I.		
COMMENT	Novel acidophil serine protease		
	Patent: JP 199032768-A 3 09-FEB-1999;		
	ONO PHARMACEUT CO LTD		
	PS Homo Sapiens		
	PN JP 1999032768-A/3		
	PD 09-FEB-1999		
	PF 16-JUL-1997 JP 1997191319		
	PR		
	PI HIROSHI KIDO,MASAHIRO INOUE		
	C12N15/09,A61K38/55,A61K39/55,A61K48/00,C07K7/64		
	PC C07K16/40,		
	PC C12N9/64,C12N15/00,A61K37/64,A61K37/64		
	CC Topology: Linear;		
	FH Key	Location/Qualifiers	
	FT CDS	10..951	
	FT sig peptide	10..87	
	FT mat peptide	133..951.	
FEATURES	Location/Qualifiers		
source	1..1082		
	/organism="Homo sapiens"		
BASE COUNT	216 a 318 c 308 g 240 t		
ORIGIN			
Query Match 96.7%; Score 1045.8; DB 6; Length 1082;			
Best Local Similarity 99.8%; Pred.No.8.9e-234;			
Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps			
Qy	31	GAGGAGGCATGGCGCGCGCGGGCGCTGCTGCTGGCGCTGCTGTGCTGGCTCGGCTGGA	90
Db	1	GAGGAGGCATGGCGCGCGCGGGCGCTGCTGCTGGCGCTGCTGTGCTGGCTCGGCTGGA	60
Qy	91	CTCAGGAAGCGGAGTCGCAAGGAGCGGGCCCCCTTATCAGGACCATTGCGGGCGACGGGTC	150
Db	61	CTCAGGAAGCGGAGTCGCAAGGAGCGGGCCCCCTTATCAGGACCATTGCGGGCGACGGGTC	120
Qy	151	ATCACGTCGCGCATCTGGGTGAGGAGGACGCCGAACCTCGGGCGTTGGCCGTGGCAGGGG	210
Db	121	ATCACGTCGCGCATCTGGGTGAGGAGGACGCCGAACCTCGGGCGTTGGCCGTGGCAGGGG	180
Qy	211	AGCCTGCGCTGTGGGATCCCACGATGCGGAGTCGAGCTCTCAGCCACCGCTGGCA	270
Db	181	AGCCTGCGCTGTGGGATCCCACGATGCGGAGTCGAGCTCTCAGCCACCGCTGGCA	240
Qy	271	CTCAGCGCGCGACTGCTTTGAAACCTATAGTAGCCTTAGTGATCCCTCCGGGTGGATG	330
Db	241	CTCAGCGCGCGACTGCTTTGAAACCTATAGTAGCCTTAGTGATCCCTCCGGGTGGATG	300
Qy	331	GTCAGTTTTGGCAGCTGACTTCATGCCATCCTTCTTGAGGCTCGAGGCTACTACACC	390
Db	301	GTCAGTTTTGGCAGCTGACTTCATGCCATCCTTCTTGAGGCTCGAGGCTACTACACC	360

QY	391	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC	450
Db	361	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC	420
QY	451	ATTGCCCTTGGTGAAGCTGTCTGCACCTGTCACTTACACTAAACACATTCAGGCCATCTGT	510
Db	421	ATTGCCCTTGGTGAAGCTGTCTGCACCTGTCACTTACACTAAACACATTCAGGCCATCTGC	480
QY	511	CTCCAGGCCCTCCACATTTTGAGCTTTGAGACCGGACAGACTGCTGGGTGACTGGCTGGGG	570
Db	481	CTCCAGGCCCTCCACATTTTGAGCTTTGAGAACCGACAGACTGCTGGGTGACTGGCTGGGG	540
QY	571	TACATCAAGAGGATGAGGCACCTGCCATCTCCACACCCCTCAGAAAGTTTCAGTCTGCC	630
Db	541	TACATCAAGAGGATGAGGCACCTGCCATCTCCACACCCCTCAGAAAGTTTCAGTCTGCC	600
QY	631	ATCATAAACACTCTATGTGCAACCACTCTTCTCAAAGTACAGTTTCGGCAAGGACATC	690
Db	601	ATCATAAACACTCTATGTGCAACCACTCTTCTCAAAGTACAGTTTCGGCAAGGACATC	660
QY	691	TTTGAGACATGGTTTTGTGCTGGCAATGCCAAGCGGGAAGATGCCCTTCGTTGGTGAC	750
Db	661	TTTGAGACATGGTTTTGTGCTGGCAATGCCAAGCGGGAAGATGCCCTTCGTTGGTGAC	720
QY	751	TCAGGTGACCCCTTGGCCTGTAAACAAGATGGACTGTGATATCAATTTGGAGTCTGTAGC	810
Db	721	TCAGGTGACCCCTTGGCCTGTAAACAAGATGGACTGTGATATCAATTTGGAGTCTGTAGC	780
QY	811	TGGGAGTGGGCTGTGTGTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCAC	870
Db	781	TGGGAGTGGGCTGTGTGTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCAC	840
QY	871	TTTCAGTGGATCCAGAAAGTGTATGGCCCAAGATGGCATGTCCCAGCCAGACCCCTCTCG	930
Db	841	TTTCAGTGGATCCAGAAAGTGTATGGCCCAAGATGGCATGTCCCAGCCAGACCCCTCTCG	900
QY	931	CCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACATCTCTGGGGCGGTTCTGAGCCTAC	990
Db	901	CCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACATCTCTGGGGCGGTTCTGAGCCTAC	960
QY	991	CTGAGCCCATGCAAGCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCTCTCTCTCTCTGTT	1050
Db	961	CTGAGCCCATGCAAGCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCTCTCTCTCTCTGTT	1020
QY	1051	TGGAATAAACACATTCAGTTGATGCCT	1079
Db	1021	TGGAATAAACACATTCAGTTGATGCCT	1049

RESULT 7	
AB031329	1082 bp mRNA linear PRI 15-JAN-2000
LOCUS	
DEFINITION	Homo sapiens esp-1 mRNA for eosinophil serine protease, complete cds.
ACCESSION	AB031329
VERSION	AB031329.1 GI:5777329
KEYWORDS	eosinophil serine protease.
SOURCE	Homo sapiens male eosinophil cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Inoue, M., Kanbe, N., Kurosawa, M. and Kido, H.
TITLE	Cloning and tissue distribution of a novel serine protease esp-1 from human eosinophils
JOURNAL	Biochem. Biophys. Res. Commun. 252 (2), 307-312 (1998)
MEDLINE	99045401
REFERENCE	2 (sites)
AUTHORS	Inoue, M., Isobe, M., Itoyama, T. and Kido, H.
TITLE	Structural analysis of esp-1 gene (PRSS 21)
JOURNAL	Biochem. Biophys. Res. Commun. 266 (2), 564-568 (1999)
MEDLINE	20068805

**REFERENCE**  
**AUTHORS** Inoue, M., Kido, H., Kanbe, N. and Kurosawa, M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (20-AUG-1999) Masahiro Inoue, Institute for Enzyme Research, Division of Enzyme Chemistry; 3 Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan (E-mail: inoue@ier.tokushima-u.ac.jp, Tel: 81-886-33-7424, Fax: 81-886-33-7425)

**FEATURES**  
 source Location/Qualifiers  
 1. .1082  
   /organism="Homo sapiens"  
   /db\_xref="taxon:9606"  
   /chromosome="16"  
   /map="16p13.3"  
   /sex="male"  
   /tissue\_type="eosinophil"  
 gene 1. .1082  
   /gene="esp-1"  
 CDS 10. .954  
   /gene="esp-1"  
   /codon\_start=1  
   /product="eosinophil serine protease"  
   /protein\_id="BAA83520.1"  
   /db\_xref="gi:577330"  
   /translation="MGARGALLALLARAGLRKPESQEAAPLSPGCGRRVTSRVG  
   GEDALGRPWQGSRLRWGLSHVCSLLSHRWALTAACFFETYSLSPPSGHWQFGQ  
   LTSMPFSWLSQAYTRYFVSNLYSPRYLGNPSYDIALVSLKLSAPYTKRHQIPCLQA  
   STFEFNRTDCWVTGWYIKDEALPSHTLOEVQVALINNSMCHFLFKYSRKDFI  
   QMDACNAGNQGKADACGDSGGLACNKNGLWYQIGVYVSWGCGGRNRPGRVYTNIS  
   HFVQIKMLQSGMSQDPSPFLPFLFLLALPLILGPV"  
 polyA\_site 1064  
   /gene="esp-1"  
   /note="18 a nucleotides"  
 BASE COUNT 216 a 318 c 308 g 240 t  
 ORIGIN  
 Query Match 96.7%; Score 1045.8; DB 9; Length 1082;  
 Best Local Similarity 99.8%; Pred. No. 8.9e-234;  
 Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Y 31 GAGGAGGCCATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGCTGGA 90  
 D 1 GAGGAGGCCATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGCTGGA 60  
 Y 91 CTCAGGAAGCCGAGTCGCGAGGAGCGCGCCCTTATCAGGACCATGCGCGCAGCGGTC 150  
 D 61 CTCAGGAAGCCGAGTCGCGAGGAGCGCGCCGCTTATCAGGACCATGCGCGCAGCGGTC 120  
 Y 151 ATCAGCTCGCGCATCTGGGTGGAGAGAGCGCGAATCGGGCGCTTGGCCGTGGCAGGG 210  
 D 121 ATCAGCTCGCGCATCTGGGTGGAGAGAGCGCGAATCGGGCGCTTGGCCGTGGCAGGG 180  
 Y 211 AGCCTCGCGCTGCGGATTCGCCAGTATCGGAGTGGAGTGGCTGCTCAGCCACCGTGGCA 270  
 D 181 AGCCTCGCGCTGCGGATTCGCCAGTATCGGAGTGGAGTGGCTGCTCAGCCACCGTGGCA 240  
 Y 271 CTCACGGCGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCGGGTGGATG 330  
 D 241 CTCACGGCGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCGGGTGGATG 300  
 Y 331 GTCCAGTTGGCCAGCTGACTTCCATGCCATCCTCTGGAGCCTGCGAGGCTTACTACAC 390  
 D 301 GTCCAGTTGGCCAGCTGACTTCCATGCCATCCTCTGGAGCCTGCGAGGCTTACTACAC 360  
 Y 391 CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTATAACATCCAGCCATGAC 450  
 D 361 CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTATAACATCCAGCCATGAC 420  
 Y 451 ATTGCCTTGGTGAAGCTGCTGCACCTGTACCTACACTAAACATCCAGCCATCTGT 510  
 D 421 ATTGCCTTGGTGAAGCTGCTGCACCTGTACCTACACTAAACATCCAGCCATCTGT 480  
 Y 511 CTCACGGCCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGTGACTGGCTGGGG 570  
 D 511



Db	721	GGACCCTGGCCCTGTAAACAAGATGACTCTGTGTATCAGATTGGAGTCGTGAGCTGGGGA	780
Qy	817	GTGGGCTGTGTCGGCCCAATCGGCCGGTGCTACACCAATATCAGCCACCACTTTGAG	876
Db	781	GTGGGCTGTGTCGGCCCAATCGGCCGGTGCTACACCAATATCAGCCACCACTTTGAG	840
Qy	877	TGGATCCAGAAGCTGATGGCCCAAGTGGCATGTCCCAGCCAGACCCCTCTCTGGCCACTA	936
Db	841	TGGATCCAGAAGCTGATGGCCCAAGTGGCATGTCCCAGCCAGACCCCTCTCTGGCCGCTA	900
Qy	937	CTCTTTTTCCTCTCTCTGGGCTCTCCACATCTCGGGCCGGTCTGAGCCTACTGAGC	996
Db	901	CTCTTTTTCCTCTCTCTGGGCTCTCCACATCTCGGGCCGGTCTGAGCCTACTGAGC	960
Qy	997	CCATGAGCCTGGGGCCATGCCAAGTCAGGCCCTGGTTCTCTCTGTCTGTTGGTAA	1056
Db	961	CCATGAGCCTGGGGCCATGCCAAGTCAGGCCCTGGTTCTCTCTGTCTGTTGGTAA	1020
Qy	1057	TAAACACATTCCAGTTGATGCCT	1079
Db	1021	TAAACACATTCCAGTTGATGCCT	1043

RESULT 9					
E21853	LOCUS				
	DEFINITION		1085 bp	DNA	linear
	ACCESSION	E21853	Novel acidophil serine protease.		
	VERSION	E21853.1			
	KEYWORDS	JP 199032768-A/5.	GI:13023724		
	SOURCE	unidentified			
	ORGANISM	unidentified			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(Bases 1 to 1085)			
	Hiroshi, K. and Masahiro, I.	Novel acidophil serine protease	Patent: JP 199032768-A 5 09-FEB-1999;	
			ONO PHARMACEUT CO LTD	
	OS	Unidentified		
	PN	JP 1993032768-A/5		
	PD	09-FEB-1999		
	PF	16-JUL-1997	JP 1997191319	

PI	HIROSHI KIDO, MASAHIRO INOUE
PC	C12N15/09, A61K38/55, A61K39/395, A61K48/00, C07K7/00,
PC	C07K16/40,
PC	C12N9/64, C12N15/00, A61K37/64, A61K37/64
CC	Strandedness: Single;
CC	Topology: Linear;
Key	Location/Qualifiers
FT	1..1085
FT	1086..1743
FT	1744..2514

FEATURES	Location/Qualifiers	/organism='Unidentified'
Source	1..1085	
		/organism='unidentified'
		/db_xref='taxon:32644'
BASE COUNT	217 a 318 c 308 q 242 t	

Query Match	95.7%	Score 1034.4	DB 6	Length 1085
Best Local	88.5%			
Still Pending				

Matches 1048;	Conservative	0;	Mismatches	1;	Indels	3;	Gaps	1;
---------------	--------------	----	------------	----	--------	----	------	----

QY	31	GAGAGGCCATGGCGCGCGCGGGGGCTCTCTGTGCGCTGTGTGCTCGGCTCGGCTGGA	90
Db	1	GAGAGGCCATGGCGCGCGCGGGGGCTCTCTGTGCGCTGTGTGCTCGGCTCGGCTGGA	60
QY	91	CTCAGGAAGCCGGAGTGCAGGAGGCGCGCCCTTATCAGGACCATCGGCCGACGGGTC	150
Db	61	CTCAGGAAGCCGGAGTGCAGGAGGCGCGCCGCTTATCAGGACCATCGGCCGACGGGTC	120
QY	151	ATCACGTCGGGCATCTGGTGGGAGGACGCCGAATCTGGGCGTTTCGGCGTGGCAGGGG	210

Db	121	ATCAGCTGCGCATCGTGGGTGGAGAGGACGCCAACTCGGGCGTTGGCGTGGCAGGGG	160
QY	211	AGCCTCGCGCTGTGGGATTCACGATGATGCGGAGTGAGCCTGCTCAGCCACCGCTGGGCA	270
Db	181	AGCCTGCGCTGTGGGATTCACGATGATGCGGAGTGAGCCTGCTCAGCCACCGCTGGGCA	240
QY	271	CTCAGCGGCGCACTGCTTTTGAACCTATAGTGACCTTAGTGACCTTCCGCGTGGATG	330
Db	241	CTCAGCGGCGCACTGCTTTTGAACCTATAGTGACCTTCCGCGTGGATG	300
QY	331	GTCAGTTTGGCCAGCTGACTTCCATGCAATCCTTCTGGAGCTCGAGGCTTACTACAC	390
Db	301	GTCCAGTTTGGCCAGCTGACTTCCATGCAATCCTTCTGGAGCTCGAGGCTTACTACAC	360
QY	391	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC	450
Db	361	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC	420
QY	451	ATTGCCCTTGGTGAAGCTGTCTGCACCTGTCACTAC --- ACTAAACACATCCAGGCCATC	507
Db	421	ATTGCCCTTGGTGAAGCTGTCTGCACCTGTCACTACTAAACACATCCAGGCCATC	480
QY	508	TGCTCTCAGGCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGG	567
Db	481	TGCTCTCAGGCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGG	540
QY	568	GGGTACATCAAGAGAGTGAAGCACTGCATCTCCACACCCCTCCAGGAAGTTCAGGTC	627
Db	541	GGGTACATCAAGAGAGTGAAGCACTGCATCTCCACACCCCTCCAGGAAGTTCAGGTC	600
QY	628	GCCATCATAAACAACTATGTGCAACCACTCTTCTCAAGTACAGTTTCCGCAAGGAC	687
Db	601	GCCATCATAAACAACTATGTGCAACCACTCTTCTCAAGTACAGTTTCCGCAAGGAC	660
QY	688	ATCTTTGGAGACATGGTTGTGTGGCAATGCCCCAAGCGGGAAGGATGCTTCGGT	747
Db	661	ATCTTTGGAGACATGGTTGTGTGGCAATGCCCCAAGCGGGAAGGATGCTTCGGT	720
QY	748	GACTCAGGTGGACCTTGGCTGTGTAACAAGATGGACTGTGGTATCAGATGGAGTGGT	807
Db	721	GACTCAGGTGGACCTTGGCTGTGTAACAAGATGGACTGTGGTATCAGATGGAGTGGT	780
QY	808	AGCTGGGAGTGGCTGTGGTTCGGCCCAATCGGCCCGGCTCTACACCAATATCAGCCAC	867
Db	781	AGCTGGGAGTGGCTGTGGTTCGGCCCAATCGGCCCGGCTCTACACCAATATCAGCCAC	840
QY	868	CACCTTGTAGTGGATCAGAGCTGATGGCCCAAGTGGCATGTCAGCCAGACCCCTCC	927
Db	841	CACCTTGTAGTGGATCAGAGCTGATGGCCCAAGTGGCATGTCAGCCAGACCCCTCC	900
QY	928	TGGCCACTACTCTTTTCCCTCTTCTGGGCTCTCCACTCTCTGGGCGGGTCTGAGCC	987
Db	901	TGGCCACTACTCTTTTCCCTCTTCTGGGCTCTCCACTCTCTGGGCGGGTCTGAGCC	960
QY	988	TACCTGAGCCCAATGACGCTGGGGCACTGCCAAGTCAAGGCCCTGGTTCTCTCTCTCT	1047
Db	961	TACCTGAGCCCAATGACGCTGGGGCACTGCCAAGTCAAGGCCCTGGTTCTCTCTCTCT	1020
QY	1048	GTTTGGTAAATAACACATTTCCAGTTGATGSCCT	1079
Db	1021	GTTTGGTAAATAACACATTTCCAGTTGATGSCCT	1052

RESULT 10  
E21854  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

E21854 Novel acidophil serine protease. 1085 bp DNA linear PAT 18-JUN-2001

E21854  
E21854.1 GI:13023725  
JP 199032768-A/6.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi.



Db	601	GCACATCAATAACAACACTATATGTGCAACACACCTCTTCTCTCAAGTACAGTTTCGCGAAGGAC	660
Qy	688	ATCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGCGGGAAGATGCCTGCTTCGGT	747
Db	661	ATCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGCGGGAAGATGCCTGCTTCGGT	720
Qy	748	GACTCAGGTGGACCCCTGGCTGTAAAGAATGGACTGTGTGATATCAGATTGGATCGTG	807
Db	721	GACTCAGGTGGACCCCTGGCTGTAAAGAATGGACTGTGTGATATCAGATTGGATCGTG	780
Qy	808	AGCTGGGAGTGGCTGTGGTGGCGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCAC	867
Db	781	AGCTGGGAGTGGCTGTGGTGGCGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCAC	840
Qy	868	CACTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGTCCAGCCAGACCCCTCC	927
Db	841	CACTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGTCCAGCCAGACCCCTCC	900
Qy	928	TGGCCACATCTCTTTTCCCTCTCTCTGGGGCTCTCCACACTCTCTGGGGCGGCTGTGAGCC	987
Db	901	TGGCCACATCTCTTTTCCCTCTCTCTGGGGCTCTCCACACTCTCTGGGGCGGCTGTGAGCC	960
Qy	988	TACCTGAGCCCATGACGAGCTGGGGCCACTGCCAAGTCAAGCCCTGTTCTCTCTGCTCTT	1047
Db	961	TACCTGAGCCCATGACGAGCTGGGGCCACTGCCAAGTCAAGCCCTGTTCTCTCTGCTCTT	1020
Qy	1048	GTTTGGTAAATAACACATTCAGTTGATGCCT	1079
Db	1021	GTTTGGTAAATAACACATTCAGTTGATGCCT	1052
RESULT 11			
AB031331			
LOCUS	AB031331	1119 bp mRNA for eosinophil serine protease 1 splicing variant, complete cds.	PRI 18-JAN-2000
DEFINITION	Homo sapiens	tesp-1	linear
ACCESSION	AB031331		
VERSION	AB031331.1	GI:6714621	
KEYWORDS	tesp-1; eosinophil serine protease 1 splicing variant.		
SOURCE	Homo sapiens	cell_line:HeLa S3	cdna to mRNA, clone_lib:Human HeLa cdna.
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Inoue M., Isobe M., Itoyama T. and Kido H.		
TITLE	Structural analysis of esp-1 gene (PRSS 21)		
JOURNAL	Biochem. Biophys. Res. Commun. 286 (2), 564-568 (1999)		
MEDLINE	20068805		
REFERENCE	2 (bases 1 to 1119)		
AUTHORS	Inoue M., Kido H. and Isobe M.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-AUG-1999) Masahiro Inoue, Institute for Enzyme Research, Division of Enzyme Chemistry; 3 Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan (E-mail: inoue@ier.tokushima-u.ac.jp, Tel: 81-886-33-7424, Fax: 81-886-33-7425)		
FEATURES	Location/Qualifiers		
source	1..1119	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="16"	
		/map="16p13.3"	
		/cell_line="HeLa S3"	
gene	1..1119	/clone_lib="Human HeLa cdna"	
		/gene="tesp-1"	
CDS	90..992	/gene="tesp-1"	
		/codon_start=1	
		/product="eosinophil serine protease 1 splicing variant"	
		/protein_id="BAA89532.1"	
		/db_xref="GI:6714622"	
		/translation="MGARGALLALLALLARAGLRKPESEAPLSPGCGRRVITSRVIG"	



QY	1	AGAGGACAGAGGGGGCGTCAGCCCGGGAGAGAGGCCATGGCGCGCGCGGGCGCTG	60
Db	51	AGGAGGACAGAGGGGGCGTCAGCCCGGGAGAGAGGCCATGGCGCGCGGGCGCTG	110
QY	61	CTGCTGGGGCTGCTGCTGGCTCGGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGCGGGC	120
Db	111	CTGCTGGGGCTGCTGCTGGCTCGGGCTGGAGTCAGGAAGCCGGAGTCGCAGGAGCGGGC	170
QY	121	CCCTTTATCAGGACCATGCGCGCAGCGGTTCATACGTCGCGCATCGTGGTGGAGAGGAC	180
Db	171	CCGTTATCAGGACCATGCGCGCAGCGGTTCATCAGTCGCGCATCGTGGTGGAGAGGAC	230
QY	181	GCCGAATCGGGGCTGGCCGTGCGAGGGAGCTCGCCTGTGGATTCACACGTATGC	240
Db	231	GCCGAATCGGGGCTGGCCGTGCGAGGGAGCTCGCCTGTGGATTCACACGTATGC	290
QY	241	GGAGTGAGCCCTGCACGCCACCGCTGGGCACTCAGCGGGCGGCATGCTTTGAAACCTAT	300
Db	291	GGAGTGAGCCCTGCACGCCACCGCTGGGCACTCAGCGGGCGGCATGCTTTGAAACCTAT	350
QY	301	AGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA	360
Db	351	AGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA	410
QY	361	TCCCTCTGGAGCCTGCGAGGCGCTACTACACCGCTTACTCGTATCGAATATCTATCTGAGC	420
Db	411	TCCCTCTGGAGCCTGCGAGGCGCTACTACACCGCTTACTCGTATCGAATATCTATCTGAGC	470
QY	421	CCTCGCTACCTGGGGAAATCACCCCTATGACATTCGCTCGTGAAGCTGCTGCACCTGTC	480
Db	471	CCTCGCTACCTGGGGAAATCACCCCTATGACATTCGCTCGTGAAGCTGCTGCACCTGTC	530
QY	481	ACCTACACTAAACACATCCAGGCCATCTGCTCCAGGCCCTCCACATTTGAGTTTGAGAAC	540
Db	531	ACCTACACTAAACACATCCAGGCCATCTGCTCCAGGCCCTCCACATTTGAGTTTGAGAAC	590
QY	541	CGGACAGACTGCTGGGTGACTGCTGGGGGTACATCAAGAGGATGAGGCATGCCATCT	600
Db	591	CGGACAGACTGCTGGGTGACTGCTGGGGGTACATCAAGAGGATGAGGCATGCCATCT	650
QY	601	CCCCACACCTCCAGGAAGTTCAAGTCGCATCAATAAACAACCTCTATGTGCAACCACTC	660
Db	651	CCCCACACCTCCAGGAAGTTCAAGTCGCATCAATAAACAACCTCTATGTGCAACCACTC	710
QY	661	TTCTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGGTTTGGCTGGCAATGCC	720
Db	711	TTCTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG-----752	752
QY	721	CAAGGCGGGAAGATGCCCTGCTTCGGTGACTCAGGTGGACCCCTTGGCCTGTAAACAAGAT	780
Db	753	-----GCTGACTCAGGTGGACCCCTTGGCCTGTAAACAAGAT788	788
QY	781	GGACTGTGTATCAGATTGGAGTCTGAGCTGGGAGTGGGCTGTGGTGGCCCCAATCGG	840
Db	789	GGACTGTGTATCAGATTGGAGTCTGAGCTGGGAGTGGGCTGTGGTGGCCCCAATCGG	848
QY	841	CCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAGAGCTGATGGCCCCAG	900
Db	849	CCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAGAGCTGATGGCCCCAG	908
QY	901	AGTGGCATGTCCACCCAGACCCCTCTCTGGCCACTACTCTTTTCCCTCTTCTCTGGGCT	960

340 GGCCAGCTGACTTCCATGCCATCCTTCGGACCTCATCGCCCTCACTACGCCGTTAATTG

RESULT	13
E21852	
LOCUS	E21852 linear DNA 945 bp Novel acidophil serine protease.
DEFINITION	Novel acidophil serine protease.
ACCESSION	E21852
VERSION	E21852.1 GI:13023723
KEYWORDS	JP 199032768-A/4.
SOURCE	unidentified.
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 945)
TITLE	Hiroshi,K. and Masahiro,I.
JOURNAL	Novel acidophil serine protease
COMMENT	Patent: JP 199032768-A 4 09-FEB-1999; ONO PHARMACEUT CO LTD OS Unidentified PN JP 199032768-A/4 PD 09-FEB-1999 PF 16-JUL-1997 JP 1997191319 PR HIROSHI KIDO,MASAHIRO INOUE PC C12N15/09,A61K38/55,A61K38/55,A61K39/395,A61K48/00,C07K7/00, PC C07K16/40, PC C12N9/64,C12N15/00,A61K37/64,A61K37/64 CC Strandedness: Single; CC Topology: Linear; FH key Location/Qualifiers FT source 1..945 /organism='Unidentified'.

[illegible]

Db 433 CCCAATGACATAGCCCTGCTGAAGTGTATCTCCAGTCACCTACAATAAAGTTCATCCAG 492  
QY 502 CCCATCTGCTCCAGGCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGTGACT 561  
Db 493 CCCATCTGCTCCAGGCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGTGACT 552  
QY 562 GGCTGGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCCCACACCTCCAGGAAGTT 621  
Db 553 GGCTGGGGGTATTTGGAGAGATGAGAGTCTGCCATCTCCACACACTCTCCAGGAAGTG 612  
QY 622 CAGTGGCCATCATTAACAACTCTATGTGCAACCACTCTTCTCAAGTACAGTTTCGC 681  
Db 613 CAGGTAGCTATTATCAACAACAGCATGTGTAAACCATATGTACAAAAAGCCAGACTTCGC 672  
QY 682 AAGGACATCTTTGGAGACATGTTTGTGCTGGCAATGCCCAAGGGGGAAGGATCCCTGC 741  
Db 673 ACGAATCTCTGGGAGACATGTTTGGCTGGCACTCTCTGAAGGTGGCAAGGATGCCCTGC 732  
QY 742 TTCGGTGACTCAGGTGGACCCCTTGCCCTGTAAACAAGATGGACTGTGGTATCAGATTGGA 801  
Db 733 TTTGGTGACTCGGAGGACCCCTTGCCCTGGCAGCCAGGATACGGTGTGGTATCAGTTGGA 792  
QY 802 GTCGTGAGCTGGGAGTGGGTGTGGTGGCCCAATCGGCCCGGTGTCTACACCAATATC 861  
Db 793 GTTGTGAGCTGGGGAATAGGCTGTGGTGGCCCAATCGGCCCTGGAGTCTATACCAACATC 852  
QY 862 AGCCACCACCTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGTCCAGCCAGAC 921  
Db 853 AGTCATCACTACAACTGGATCCAGTCAACCATGATCCGAATGGGCTGCTCAGGCCCTGAC 912  
QY 922 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGGCCGGTC 981  
Db 913 CCAGTCCCTTTGCTACTGTTTCTTACTCTGGCCTGGGCTTCTCTTGTCTGAGGCCCTGCC 972  
QY 982 TGAGCCTACCTGAGCCCATGAGCCTGGGGCCACTGCCAAGTACAGGCCCTGTTCTCTTC 1041  
Db 973 TGAGCCACACGTTAGCTACACCTGTG-----AGGTCAGGGTGTGTCTCTTTTG 1023  
QY 1042 TGTCTTGTGTTGTAATAACACATT 1066  
Db 1024 TATCTTGTGCTAATAAACCTGTT 1048

Search completed: May 10, 2003, 04:20:48  
Job time : 3364 secs

**THIS PAGE BLANK (USPTO)**

09/787, 844

SYSTEM:OS - DIALOG OneSearch

File 5:Biosis Previews(R) 1969-2003/Oct W1

(c) 2003 BIOSIS

File 73:EMBASE 1974-2003/Oct W1

(c) 2003 Elsevier Science B.V.

File 155:MEDLINE(R) 1966-2003/Oct W1

(c) format only 2003 The Dialog Corp.

**\*File 155: Medline has been reloaded and accession numbers have changed. Please see HELP NEWS 155.**

File 159:Cancerlit 1975-2002/Oct

(c) format only 2002 Dialog Corporation

**\*File 159: Cancerlit ceases updating with immediate effect.**

Please see HELP NEWS.

Set	Items	Description
S1	8959	CGS
S2	588	S1 AND CANCER?
S3	51	S2 AND (GYNECOL? OR TESTI?)
S4	44	RD (unique items)

5/9/1 (Item 1 from file: 5)

DIALOG(R)File 5:Biosis Previews(R)

(c) 2003 BIOSIS. All rts. reserv.

11134686 BIOSIS NO.: 199799755831

**A novel in vitro and in vivo breast cancer model for testing inhibitors of estrogen biosynthesis and its action using mammary tumor cells with an activated int-5/aromatase gene.**

AUTHOR: Tekmal Rajeshwar R(a); Durgam Vijayender R

AUTHOR ADDRESS: (a)Dep. Gynecol. Obstet., Winship Cancer Cent., Emory Univ. Sch. Med., 4219 WMB, 1639 Pierce Drive,\*\*USA

JOURNAL: Cancer Letters 118 (1):p21-28 1997

ISSN: 0304-3835

RECORD TYPE: Abstract

LANGUAGE: English

ABSTRACT: We recently showed that the cellular gene int-5/aromatase in BALB/c mammary alveolar hyperplastic nodule (D2 HAN/D2 tumor cells) is activated as a result of mouse mammary tumor virus integration within the 3' untranslated region of the aromatase gene. In the present study, we evaluated the effect of various aromatase inhibitors on androstenedione-mediated tumor cell growth. Also, we compared the effect of the non-steroidal aromatase inhibitor (CGS 16949A) on the inhibition of tumor growth. Our results show that D2 tumor cells respond well to various aromatase inhibitors and antiestrogens. We examined the usefulness of this model by using D2 tumor cells to simulate postmenopausal breast cancer employing both in vitro cell culture and in vivo ovariectomized (OVX) nude mouse. Unlike DMBA-induced tumors or other models, D2 tumor cells form very rapid tumors within a few days in intact mice or OVX nude mice with androstenedione supplementation and respond well to an aromatase inhibitor. This model with its known mechanism of aromatase activation should be useful for studying the role of intra-tumoral estrogen in mammary cancer, for evaluating the effects of aromatase inhibitors and antiestrogens, and for comparing breast cancer treatments.

5/9/2 (Item 2 from file: 5)

DIALOG(R)File 5:Biosis Previews(R)

(c) 2003 BIOSIS. All rts. reserv.

09457072 BIOSIS NO.: 199497465442

**Aromatase inhibitors - mechanisms for non-steroidal inhibitors.**

AUTHOR: Vanden Bossche Hugo(a); Moereels Henri; Koymans Luc M H

AUTHOR ADDRESS: (a)Dep. Comparative Biochem., Janssen Research Foundation, B-2340 Beerse\*\*Belgium

JOURNAL: Breast Cancer Research and Treatment 30 (1):p43-55 1994

ISSN: 0167-6806

DOCUMENT TYPE: Literature Review

RECORD TYPE: Abstract

LANGUAGE: English

ABSTRACT: The conversion of androgens to estrogens occurs in a variety of

**THIS PAGE BLANK (USPTO)**

cells and tissues, such as ovarian granulosa and **testicular** cells, placenta, adipose tissue, and various sites of the brain. The extragonadal synthesis of estrogens has great pathophysiological importance. Estrogens produced by, for example, adipose tissue have a role in the pathogenesis of certain forms of breast **cancer** and endometrial adenocarcinoma. The biosynthesis of estrogens is catalyzed by the aromatase, an enzyme localized in the endoplasmic reticulum that consists of two components: a cytochrome P450 (P450 Arom, P450 19 product of the CYP19 gene) and the NADPH cytochrome P450 reductase. The alignment of the amino acid sequences of human P450 19 with other mammalian P450s shows little sequence similarity, which indicates not only that P450 19 is a unique form of the P450 superfamily but also that the aromatase may be a good target, for the development of selective P450 inhibitors. Aminoglutethimide (AG) is the pioneer drug of the reversible competitive nonsteroidal aromatase inhibitors. Since AG is a nonspecific aromatase inhibitor and presents some problems with tolerability, a number of structural analogues have been synthesized. For example, rogletimide is slightly less potent than AG but has the advantage of not inhibiting the cholesterol side-chain cleavage and is devoid of sedative action. Elongation of the ethyl substituent of AG and rogletimide leads to an increase in aromatase inhibition. Further studies led to the discovery of a new generation of much more potent aromatase inhibitors. An example is fadrozole. However, although fadrozole is a poor inhibitor of the cholesterol side-chain cleavage, it suppresses aldosterone release by ACTH-stimulated human adrenocortical cells. More selective aromatase inhibitors are the triazole derivatives. Examples are **CGS** 20267, **CGS** 47645, R 76 713, and ICI D1033. R 76 713's aromatase inhibitory effect is largely due to its (+)-S-enantiomer, vorozole. Computer modeling studies of the interaction of vorozole with part of the "I-helix" of P450 19 suggest that the chlorine-substituted phenyl ring of vorozole interacts with the gamma-carbonyl group of Glu-302. Thr-310, which corresponds to the highly conserved Thr-252 in P450 101, interacts with vorozole's triazole ring, and the 1-methyl-benzotriazole moiety binds near Asp-309.

REGISTRY NUMBERS: 9039-48-9: AROMATASE; 112809-51-5: **CGS** 20267;  
118949-22-7: R 76713; 129731-10-8: VOROZOLE; 121840-95-7: ROGLETIMIDE



**THIS PAGE BLANK (USPTO)**

*Search notes***WEST****The Contents of Case 09787844**

---

Qnum	Query	DB Name	Thesaurus	Operator	Plural
Q1	cancer specific antigen	USPT	ASSIGNEE	ADJ	YES
Q2	Q1 and ovarian cancer	USPT	ASSIGNEE	ADJ	YES
Q3	cancer specific antigen.clm	USPT	ASSIGNEE	ADJ	YES
Q4	her2.clm	USPT	ASSIGNEE	ADJ	YES
Q5	her2.clm.	USPT	ASSIGNEE	ADJ	YES
Q6	cancer specific antigen.clm.	USPT	ASSIGNEE	ADJ	YES

---

[Run Case](#)[Update Case](#)[Cancel](#)[Help](#)[Main Menu](#)[Logout](#)

**THIS PAGE BLANK (USPTO)**